REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	May 29, 1997	Final Technical Report 1993-1997

4. TITLE AND SUBTITLE
Collisionless Dynamics of the Magnetosphere

5. FUNDING NUMBERS
F49620-93-1-0071

6. AUTHOR(S)
Michael E. Mauel, Columbia University

Amitava Bhattacharjee, University of Iowa

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Columbia University
Office of Projects&Grants
500 West 120th St., Rm351 Eng Ter.
New York, NY 10027

University of Iowa
Dept.ofPhysics&Astronomy
211 Van Allen Hall
Iowa City, Iowa 52242-1410

8. PERFORMING ORGANIZATION REPORT NUMBER
N/A

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Office of Scientific Research AFOSR/NM
110 Duncan Ave., Suite B115
Bolling Air Force BAse, D.C. 20332-001

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release;
distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

[Abstract for AFOSR Award Number F49620-93-1-0071]

This research program contributed to the development of a physics-based model for substorm formation and the resulting intense modulation of energetic particle flux observed in the Earth's magnetosphere. Specifically, this research (1) demonstrated the relationship between the spectrum and intensity of low-frequency fluctuations and the collisionless transport of energetic trapped particles, (2) verified the applicability of the guiding-center drift Hamiltonian as the foundation for numerical simulation of driven particle flows, (3) derived a unified analytical theory of the electron and ion tearing instability in the presence of all three components of the tail magnetic field in order to identify a trigger for substorms, and (4) investigated the effect of small but finite resistivity on the dynamics of current-sheet formation in the solar corona. All topics originally proposed for study have been addressed by our studies and published within archival journals. In addition, our research uncovered entirely phenomena associated with complex nonlinear interactions within a driven collisionless plasma. These topics are the subject of new and separately funded research activities.

14. SUBJECT TERMS

DTIC QUALITY INSPECTED 4

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT

AFOSR-TR
97-0616

Final Technical Report
AFOSR Award Number F49620-93-1-0071

TITLE OF RESEARCH PROGRAM: Collisionless Dynamics of the Magnetosphere

SUBMITTED BY: Michael E. Mauel, Professor
Department of Applied Physics
Columbia University, New York, NY 10027

Amitava Bhattacharjee, Professor
Department of Physics and Astronomy
University of Iowa, Iowa City, 52242

Our research program (1) demonstrated the relationship between the spectrum and intensity of low-frequency fluctuations and the collisionless transport of energetic trapped particles, (2) verified the applicability of the guiding-center drift Hamiltonian as the foundation for numerical simulation of driven particle flows, (3) derived a unified analytical theory of the electron and ion tearing instability in the presence of all three components of the tail magnetic field in order to identify a trigger for substorms, and (4) investigated the effect of small but finite resistivity on the dynamics of current-sheet formation in the solar corona. All topics originally proposed for study have been addressed by our studies and published within archival journals. In addition, our research uncovered entirely new topics of investigation which are the subject of new and separately funded research activities.

In the following, we present a brief summary of previous work and attach reprints of our publications. This summary is separated into (1) a report of the results from the laboratory investigation of collisionless energetic particle transport induced by fluctuations, and (2) a report of the subcontract with the University of Iowa summarizing the conclusions of new theories applied to substorm modeling and current sheet formation.

1. Laboratory Investigation of Collisionless Energetic Particle Transport

In this part of our research, we summarize experimental and computational investigations of the effects of drift-resonant fluctuations, chaotic transport of trapped,

19971204 128

energetic electrons, threshold for global chaotic transport, and particle simulation using the guiding center drift Hamiltonian and permitting a detailed comparison between the chaotic evolution of phase-space and laboratory observations of intense collisionless radial transport. This research resulted in three referred publications, several invited lectures, and one doctoral dissertation.

We divided our research into three main parts: (1) formation and observation of an artificial radiation belt, (2) measurement of collisionless fluctuation-induced transport, and (3) modeling and interpretation of the experimental results. *The major result from this research is the detailed demonstration of the relationship between collisionless radial transport and the spectral intensity of drift-resonant fluctuations.*

FORMATION AND OBSERVATION OF AN ARTIFICIAL RADIATION BELT

Our first laboratory tasks were (1) to complete construction of the Colllisionless Terrella Experiment (CTX) and (2) to produce and characterize an "artificial radiation belt" of energetic electrons.

The CTX represented the first attempt to produce a collisionless, high-energy plasma confined by a laboratory dipole magnetic field, and optimization of the energetic electron production technique and development of new plasma diagnostics were required. We were succeeded to produce and sustain a non-Maxwellian population of energetic electrons with microwave heating [1,2] characterized by a "power-law" energy spectrum having magnetospheric relevance. A relatively low background neutral gas density gave the energetic electrons a very long mean-time between collisions, $\tau_{col} > 10^3/\omega_d$, and their motion was essentially adiabatic in the absence of resonant fluctuations. Since plasmas were created using electron cyclotron resonance heating (ECRH), the heating was localized at a particular radius. An intense ring of energetic electrons would form, and we referred to this localized population of energetic electrons as an "artificial radiation belt."

Using a variety of movable probes, we observed Hot Electron Interchange (HEI) instabilities. In a dipole-confined plasma, the HEI instability appeared with a complex and evolving spectrum of multi-mode waves resonating with the precessional drifts of the energetic electrons. The waves are excited when the radial density gradient of the artificial radiation belt steepens during microwave heating. These waves served as laboratory "test waves" for our investigation of collisionless chaotic transport of trapped particles in dipolar magnetic fields. In a manner of speaking, they are the analogs to the externally driven and the internally excited global variations of the Earth's geomagnetic and electric fields.

MEASUREMENT OF COLLISIONLESS FLUCTUATION-INDUCED TRANSPORT

Our next significant achievement was the detailed measurement of the electrostatic fluctuations and induced electron transport. Using high-impedance, high-frequency electrostatic probes and miniature, gridded particle energy analyzers, we were able to measure *both* wave-spectra leading to global chaotic transport and wave-spectra leading to saturated, isolated bands of stochastic drifts. We found the wave-amplitude alone to be a poor indicator of chaos. Both the spectral content and the wave amplitude were needed to determine the onset and magnitude of wave-induced transport. Energetic electron transport was observed only when the wave-spectra was consistent with our predictions of global chaos. The fluctuations resulted in more than an order-of-magnitude increase in detected particle flux. Additionally, we discovered a strong temporal modulation of the energetic electron flux induced by the fluctuations.

MODELING AND INTERPRETATION OF THE EXPERIMENTAL RESULTS

The detailed experimental results obtained from CTX enable the first global simulation of collisionless radial transport of energetic particles induced by drift-resonant fluctuations having complex spectral content. These results are significant since they establish a physics basis for similar modeling of energetic particle transport induced by electric and magnetic fluctuations in the Earth's magnetosphere.

Our description of particle dynamics driven by large scale fluctuations is based on a guiding center Hamiltonian formalism. The guiding center Hamiltonian facilitates comparison between processes observed in the laboratory with those observed in the magnetosphere. Since we were able to observe energetic electron transport only when the wave-spectra was consistent with our predictions of global chaos, we have been able interpret our laboratory results in terms of fundamental properties of radial transport in dipole magnetic fields. By conducting relatively simple particle simulations based upon the guiding-center drift Hamiltonian and using the spectral intensity of the fluctuations observed experimentally, we successfully modeled the observed transport [2] and developed detailed understanding of the temporal modulations of the local energetic particle flux and of the reasons for the presence or absence of transport as the spectral content broadens [3]. The temporal modulations of the energetic particle flux measured at a localized particle detector is analogous to drift echoes observed following a impulsive events in the magnetosphere. Drift resonant fluctuations which do not lead to global

chaos are analogous to various geomagnetic pulsations which modulate particle phase-space without causing transport.

Although we have been able to compare theoretically the observed transport rates with quasilinear predictions, we have not yet measured global transport rates or the profile evolution of the energetic particle flux. New diagnostic arrays have been installed in order to measure profile evolution of the x-ray bremsstrahlung emission, neutral light emission, and polar precipitation rates [4]. These measurements are critically important to our goals (1) of examining radial chaotic transport which preserves the first two adiabatic invariants, μ, J, (2) of observing the natural tendency for phase-space profiles to evolve to "stationary" profiles, $\partial F / \partial \psi \sim 0$, during chaotic transport, and (3) of understanding the relationship of these profiles to dipole plasma stability criteria. These measurements are used as the basis of a new diagnostic array presently being installed.

The observations from CTX have also excited interest in the wider plasma physics community, especially scientists investigating plasma chaos and nonlinear behaviors. The processes represented by adiabatic radial transport (and illustrated by storm-time inward diffusion of particles and energy) is being studied as an explanation of the anomalous pinch effect seen in toroidal magnetic plasma devices. The slow, rising tones, or "chirping", seen in CTX appears to be ideally suited to investigate time-evolving, nonlinear wave-particle resonances. Under certain conditions, frequency modulation leads to enhanced transport due to "dynamic autoresonance", unlimited acceleration, and "bursting" and "blow-up". Particularly relevant to CTX observations is the possible transport of phase-space "holes" or "clumps" generated by slowly evolving wave-particle resonances. At this time, we are not certain of the importance of these related phenomena to processes within the terrestrial magnetosphere. However, the physics underlying this phenomena is generally applicable to all instability processes near marginal criticality, and future observations in space may still discover events involving wave-particle resonant trapping—provided that multi-point wave measurements are deployed. Indeed, Chen and co-workers, observe "resonance broadening" effects in the storm-time ring current, and nonlinear wave-trapping may be the explanation for energy-dependent enhancements of the radial diffusion coefficient.

The CTX device has enabled unique and significant tests of (1) the application of guiding center theory to large-scale transport of energetic plasma within a dipole magnetic field, and (2) recent developments in nonlinear wave-particle dynamics. These include:

1. Demonstrated the utility of the guiding-center drift Hamiltonian to describe chaotic particle dynamics during intense drift-resonant fluctuations.

2. Demonstrated the relationship between the spectral-content of the fluctuations and the existence of global transport. Chaotic transport rates depend on both the amplitude and spectrum of the fluctuations.

3. Explored a unique laboratory facility for the investigation of complex wave-particle dynamics of energetic trapped particles in a dipole magnetic field. CTX allows for the investigation of frequency chirped waves and nonlinear amplification processes due to the inward transport of phase-space holes.

Several questions remain to be answered, and these are subjects of ongoing research.

2. BIBLIOGRAPHY REFERRED TO IN PART 1

2.1 Articles

1. "Observation of Drift-Resonant Fluctuations in the Collisionless Terrella Experiment," M. E. Mauel and H. P. Warren, in Physics of Space Plasmas (1993), SPI Conference Proceedings and Reprint Series, No. 12, T. Chang, G. B. Crew, and J. R. Jasperse, eds. (Scientific Publishers, Cambridge, MA, 1994).

2. "Observation of Chaotic Particle Transport Induced by Drift-Resonant Fluctuations in a Magnetic Dipole Field," H. P. Warren and M. E. Mauel, *Phys. Rev. Letters* **74** (1995) 1351.

3. "Wave-Induced Chaotic Radial Transport of Energetic Electrons in a Laboratory Terrella Experiment", H. P. Warren and M. E. Mauel, *Phys. of Plasmas* **2** (1995) 4185.

4. "Observation of wave-induced chaotic radial transport in a laboratory terrella experiment," H. P. Warren, M. E. Mauel, D. Brennan, and S. Taromina, *Phys. of Plasmas* **3** (1996) 2143.

2.2 Invited Lectures and Conferences

5. "Observation of Drift-Resonant Fluctuations in the Collisionless Terrella Experiment," 2nd Intern. Workshop of Space and Laboratory Plasmas, Banff, Canada, July, 1993, Cambridge Workshop on Space Plasmas, M.I.T., August, 1993, Phillips Laboratory, Hanscom AFB, February, 1994, and University of New Hampshire, April, 1994, U.C. Irvine, September, 1994, U.C. San Diego, December, 1994, Cornell University, February, 1995.

6. "Observation of wave-induced chaotic radial transport in a laboratory terrella experiment," H. P. Warren, 3nd Intern. Workshop of Space and Laboratory Plasmas, Banff, Canada, July, 1995.

7. "Progress in Fusion and Laboratory Experiments in the Development of Plasma Physics," M. E. Mauel, 3nd Intern. Workshop of Space and Laboratory Plasmas, Banff, Canada, July, 1995.

8. "Observation of wave-induced chaotic radial transport in a laboratory terrella experiment," H. P. Warren, Division of Plasma Physics, APS (1995).

2.3 Doctoral Dissertations

9. Warren, H. P., *Observation of Chaotic Particle Transport Driven by Drift-Resonant Fluctuations in the Collisionless Terrella Experiment*, Ph.D. Dissertation, Columbia University, 1995.

OBSERVATION OF DRIFT-RESONANT FLUCTUATIONS IN THE COLLISIONLESS TERRELLA EXPERIMENT

M. E. Mauel and H. P. Warren
Department of Applied Physics
Columbia University, New York, NY 10027 U.S.A.

ABSTRACT

We report the first observation of drift-resonant instabilities of a laboratory plasma confined with a dipole magnetic field. The instabilities appear when a sufficiently energetic belt of electrons is created with microwave heating. During heating, intense, quasi-periodic "bursts" of drift-resonant waves are excited and induce rapid electron transport. Immediately after the heating is switched off, the instability's characteristics change and relatively long-lasting rising tones are observed. In both cases, broad-banded multiple modes are associated with collisionless transport as required for chaotic transport in a dipole magnetic field. Absolute measurement of the wave spectrum at the onset of the transport allows direct test of the theory of transport induced by intense waves.

I. INTRODUCTION

The "Collisionless Terrella Experiment", or CTX, is a laboratory experiment built at Columbia University in order to directly observe the collisionless radial transport of plasma trapped within a dipole magnetic field. Motivated by Hasegawa [1,2], CTX aims to characterize the wave intensity and spectrum required for the onset of chaotic transport and to observe the time evolution of plasma profiles undergoing rapid transport.

Models of the energetic trapped-particle diffusion within planetary magnetospheres are fundamental to our understanding of plasma dynamics [3]. Nakada and Mead [4] described radiation belt diffusion in terms of the observed frequency spectrum of magnetic fluctuations associated with sudden impulses and sudden commencements. Birmingham [5] described radial diffusion induced by fluctuating convection electric fields. In the classic paper by Farley, Tomassian and Walt [6] the observed radial distribution of high-energy protons in the inner radiation belt was understood only when inward and outward diffusion induced by drift-resonant magnetic fluctuations were included. Tomasian, *et al.* [7] noticed a strong correlation between the inner zone electron fluxes with E field fluctuations measured following substorms. Each of these models assume the preservation of the first two adiabatic invariants, μ, and J, during the chaotic dynamics of the third adiabatic, ψ, (which is proportional to

the magnetic flux enclosed by a drift surface.) Warren, *et al.*, [8] recently demonstrated the validity of this assumption, provided a sufficiently large separation exists between the particle's bounce, ω_b, and drift, ω_d, frequencies. Under these conditions, the evolution of the plasma distribution function, $F(\mu, J, \psi, t)$ can often be diffusive

$$\frac{\partial F}{\partial t} = S - \frac{F}{\tau} + \frac{\partial}{\partial \psi} D \frac{\partial F}{\partial \psi} \tag{1}$$

where S is a net particle source, τ is polar precipitation time, and $N(\psi, t) = \int d\mu\, dJ\, F(\mu, J, \psi, t)$ is the total number of particles within a tube of unit flux, $\delta\psi$. As shown by Chan, *et al.* [9], the diffusion coefficient can be written in the quasilinear limit as

$$D = \frac{\pi}{2} \sum_{\substack{m \neq 0 \\ \omega}} \delta(\omega - m\omega_d) \left[m^2 \left| \delta\Phi \right|^2 + \omega^2 \left| \delta A_\varphi \right|^2 \right] \tag{2}$$

where $|\delta\Phi|^2$ is the intensity of electric fluctuations, $|\delta A_\phi|^2$ is related to the intensity of magnetic fluctuations, (ω, m) are the fluctuation's frequency and azimuthal mode number, and $(\pi/2)\ \delta(\omega - m\,\omega_d)$ is the quasilinear resonance function. For example, the energy-independent diffusion coefficient used successfully in Ref. 6 results from the spectrum of the $m = 1$ component of the magnetic fluctuations induced by Si and Sc events having $|\delta A_\phi|^2 \propto \omega^{-2} L^4$. (Here, L refers to the McIlwain radial coordinate.)

Observations of Pc5 drift-resonant pulsations [10] having higher azimuthal mode numbers, $|m| > 30$, have motivated studies of drift-bounce particle resonances [9]. The interactions between these pulsations and barely trapped O^+ ring-current ions have been proposed to explain the significant loss of 5- to 17-keV ions during the recovery phase of geomagnetic storms [11].

The implications of Eqs. 1 and 2 need emphasis. For sufficiently intense wave-particle interactions, radial diffusion dominates sources and polar loses, (*i.e.* S and $1/\tau$), and F evolves until $\partial F/\partial\psi \approx 0$. $N(\psi)$ is constant, and the local plasma density varies as $n \sim L^{-4}$. Similarly, the pressure profile scales as $p \sim L^{-20/3}$, corresponding also to marginal stability for interchange and ballooning modes [12]. When p varies more steeply than $L^{-20/3}$, $\partial F/\partial\psi > 0$, and internally-excited drift-resonant instabilities can drive hot plasma outward. When p varies less steeply than $L^{-20/3}$, $\partial F/\partial\psi < 0$, and externally-excited fluctuations can lead to inward diffusion—further steepening the plasma's radial density gradient, dn/dL.

The applicability of Eqs. 1 and 2 is being directly tested in the CTX experiment for a variety of electrostatic fluctuation spectra. The fluctuations in CTX do not *simulate* the fluctuations observed in the magnetosphere. Instead, they enable investigation of the fundamental *process* of collisionless radial transport in a controlled laboratory experiment. In Section II, we briefly describe the experimental apparatus and the formation of an energetic electron

belt with microwave heating. In Section III, we describe observations of drift-resonant instabilities exciting rapid radial transport of the energetic electrons. In Section IV, we discuss the implication of these observations to trapped-particle dynamics in planetary magnetospheres. We note that collisionless radial transport is observed only when wave-particle interactions are sufficiently intense and broad-banded.

II. DESCRIPTION OF EXPERIMENT

The CTX experiment consists of (1) a large nearly spherical vacuum chamber approximately 1.6 m in diameter, (2) an internal dipole magnetic suspended along the axis of one magnetic pole by a stainless-steel tube carrying electrical power and water cooling, and (3) a microwave plasma source which can be used to make either a "plasma stream" or a source of microwaves for direct heating of electrons, magnetically trapped by the dipole. External electromagnets can be used to apply an axial "bias" field and form a closed magnetic cavity for the dipole. For the experiments described here, we have operated the microwave source to maximize the production of an "artificial radiation belt" consisting of trapped electrons with energies between 1 and 40 keV. These electrons have a collisional mean-free-path longer than 1,000 drift orbits about the equator, and they provide the ideal laboratory medium with which to study radial transport in dipole geometry.

Fig. 1 shows the time history of a typical CTX discharge and the production of an energetic electron belt. The discharge duration was

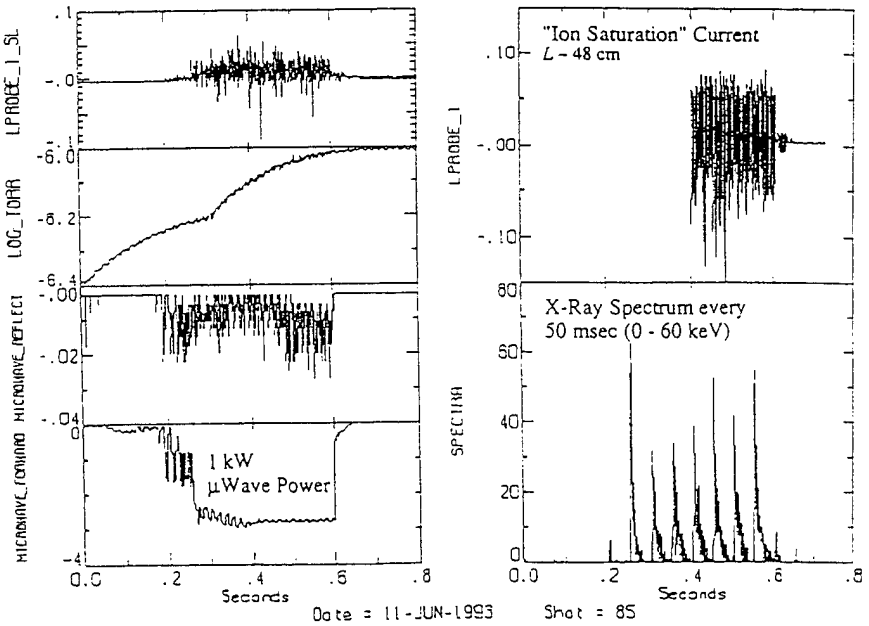

Figure 1. Several signals illustrating the time history of a typical CTX discharge containing an energetic electron belt.

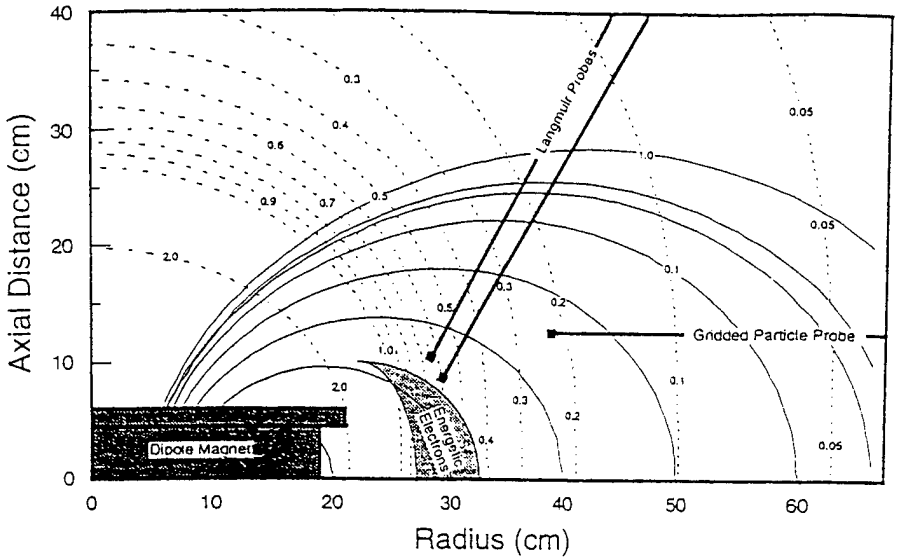

Figure 2. Magnetic geometry of the dipole field showing belt location, probe locations, magnetic field strength (dotted lines), and field lines (solid lines).

programmed to last approximately 0.5 sec, and the several signals represent the time history of the ion-saturation current to a Langmuir probe, the hydrogen gas pressure, the forward and reflected microwave power, and the x-ray spectra as recorded by a krypton proportional counter. The experiment is fully computer controlled, and the gas pressure, pulse length, and heating power can be programmed independently for each discharge. For discharges similar to that shown in Fig. 1, intense fluctuations are observed both during the microwave heating and during the "afterglow" when the heating power has been switched-off.

Multiple Langmuir probes are used to detect the drift-resonant waves and measure the background plasma parameters. For moderate and low hydrogen gas fill pressures, $P < 2 \times 10^{-6}$ Torr, the plasma density is less than 10^{10} cm^{-3} and the background electron temperature is approximately 100 eV. The proportional x-ray detector is used to estimate the energy distribution of the energetic electrons. As shown in Fig. 2, the energy distribution is non-Maxwellian and resembles the power-law distributions characteristic of wave heating (in this case cyclotron heating) [13]. The intensity of the electron belt depends sensitively on the background hydrogen pressure, allowing experimental control of electron energy. Fig. 3 shows a quadrant of a cross-sectional view of the magnetic geometry of the experiment illustrating the location of the probes and energetic electron belt. Direct contact between the belt and a probe reduces the intensity of the energetic electrons and this technique is used to determine the belt's location.

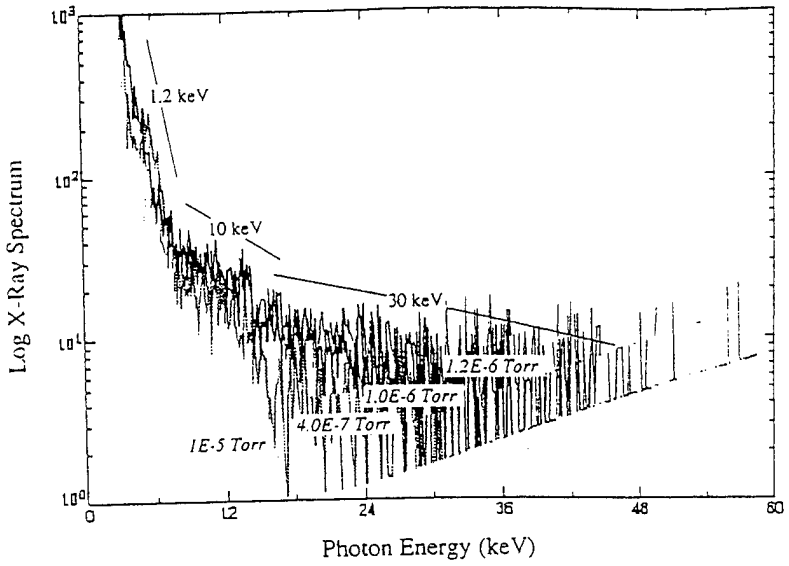

Figure 3. Measured x-ray spectra emitted from the energetic electron belt showing the dependence of electron intensity on hydrogen fill pressure.

III. OBSERVATION OF DRIFT-RESONANT INSTABILITIES

Fig. 4 shows the probe fluctuations on a faster, 10 msec, time-scale. During the microwave heating, quasi-periodic "bursts" are observed; whereas, during the afterglow, the wave amplitude evolves more gradually. As shown, in Fig. 5, the quasi-periodic bursts correspond to (a) large negative reductions in the plasma floating potential and (b) rapid radial transport of electrons from the belt region.

By using very high-speed data recorders, we have identified the drift-resonant instabilities inducing this rapid electron transport. Fig. 5 shows a frequency spectrogram of the instabilities occurring both during the microwave heating and during the afterglow. During the heating, the quasi-periodic pulsations consist of relatively wide-band signals ranging from 0.1 MHz $\leq f <$ 2 MHz. In contrast, the instabilities observed during the afterglow consist of a multi-mode collection of relatively coherent rising tones. Examination of the quasi-periodic bursts during heating also show a collection of rising tones. The difference in the frequency spectra can be linked to the energy of the electron belt as measured with the x-ray proportional counter. When the average belt energy is relatively low, $\langle E \rangle \sim$ 1-10 keV, (such as found during the heating), the wave frequency is also relatively low. During the afterglow, when the cooler electrons scattering into the terrella's polar regions, $\langle E \rangle$ increases, and the frequency of drift-resonant instabilities also increase. The rate of rise of the frequency is linked to the average energy of the trapped electron belt.

Multiple probes are used to determine the azimuthal and radial structure of the drift-resonant waves. Fig. 6 shows the magnitude and phase of the correlation of two probes located on the same flux-surface. The two probes indicate that the waves propagate in the electron drift direction, and the wave spectrum consists of multiple azimuthal mode numbers, m, as well as multiple frequencies.

These fluctuations only occur in the presence of the energetic electron belt, and we believe them to be related to drift-resonant hot electron interchange instabilities (HEI) [13,14]. During microwave heating, the electron energy distribution contains a large fraction of "warm" electrons with energies ranging from 1 to 10 keV. The instability frequency is less than the ion cyclotron frequency, ω_{ci}, and the HEI instability is predicted to have rapid growth when profile of fractional density of energetic electrons (denoted by $\alpha(\psi)$) has a sufficiently steep gradient,

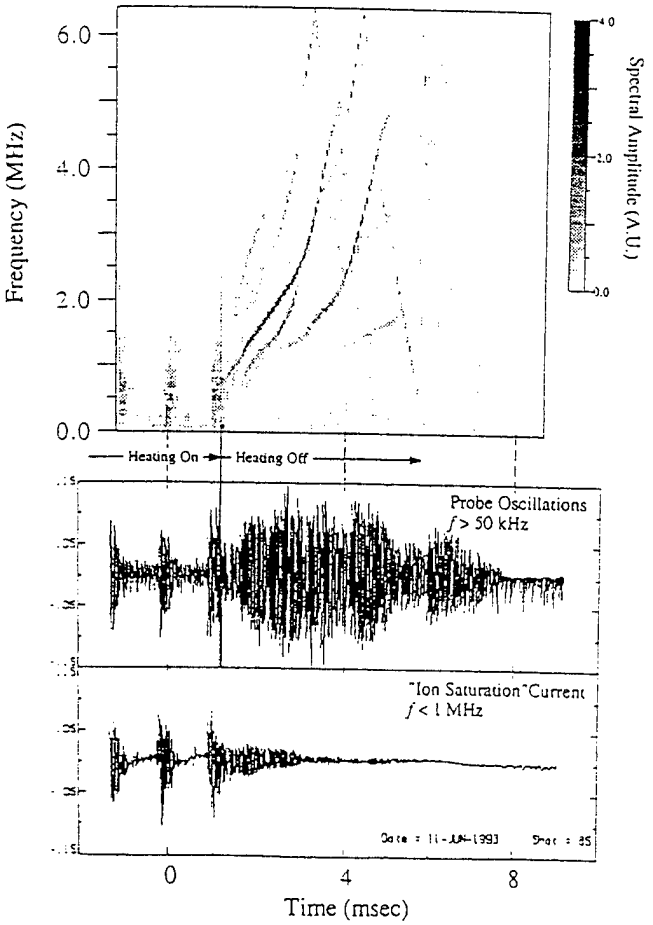

Figure 4. Frequency spectrograms showing both the quasiperiodic "bursts" during electron heating and the multiple rising tones observed during the afterglow.

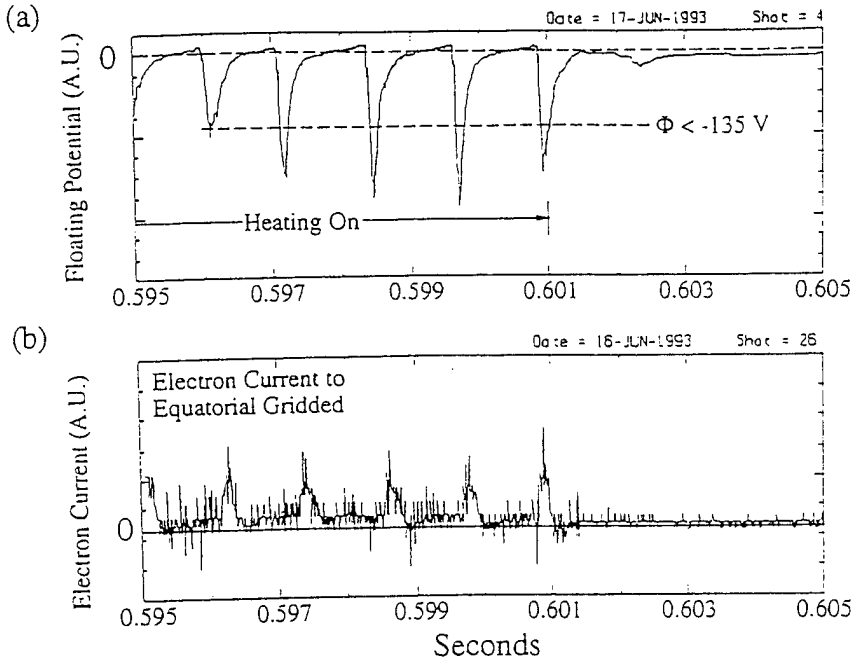

Figure 5. Rapid radial electron transport observed by (a) sharp drop of floating potential and (b) pulses of electron current observed with gridded particle detector.

$$\psi \frac{d\alpha}{d\psi} > \frac{1}{4} \frac{m\langle \omega_d \rangle}{\overline{\omega_{ci}}} \tag{3}$$

where $\overline{\omega_{ci}}$ is related to the flux-tube average of ω_{ci}. In the hot electron afterglow, the "warm" electrons are no longer reheated by the microwaves, and the average belt energy increases beyond 10 keV. The observed rising tones now represent the higher-frequency limit for HEI when $\omega \approx m\langle \omega_d \rangle \gg \omega_{ci}$ and the ions are approximately unmagnetized. In both limits, the energy to excite wave growth stems from the adiabatic cooling accompanying the radial transport induced by drift-resonant wave-particle interactions. The resulting diffusion (for broad-banded spectra) and wave-trapping (for larger more coherent spectra) ultimately leads to wave damping or saturation.

IV. DISCUSSION

We report observations of rapid collisionless radial transport induced by an intense spectrum of drift-resonant fluctuations. These results have implications for understanding the basic process of collisionless radial transport in a dipole magnetic field. First, transport in a dipole magnetic field seems to require multiple modes. This is a unique property of the dipole magnetic geometry resulting from the strong radial dependence of a particle's azimuthal drift frequency, $\omega_d \propto 1/L^2$, where L is the equatorial radius of a flux surface.

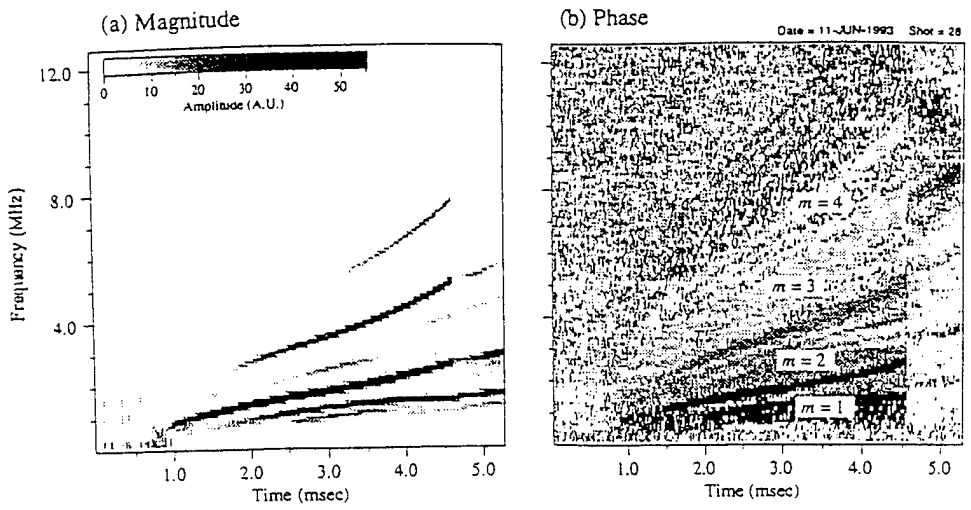

Figure 6. Magnitude and phase of the correlation between two probes positioned at the same flux surface showing the excitation of multiple modes.

Global transport can only occur with multiple modes since a particle does not remain correlated with a single wave as it diffuses radially. Secondly, the radial transport rate is fast. In CTX, large electron bursts lasting only 10's of drift periods are observed to cause significant transport. Finally, drift-resonant instabilities lead to transport causing self-stabilization and continuous, quasi-periodic bursting. This implies that the transport-inducing waves also decrease the radial pressure gradient of the electron belt—but do not destroy electron confinement altogether. For a dipole magnetic field, the marginally stable pressure profile scales like $p \propto L^{-20/3}$, and it is likely the "bursts" of drift-resonant instabilities maintain this profile in opposition to the tendency of the microwave heating to steepen the belt pressure gradient.

We have recently completed experiments quantifying the wave amplitude, drift-island size, and the temporal details of the energetic radial electron flux. These results will be reported elsewhere. Future experiments will use electron cyclotron emission to measure the evolution of the pressure profile and externally-imposed magnetic impulses to generate radial transport.

ACKNOWLEDGMENTS

The authors are grateful to the Air Force Office of Scientific Research (Grant No. F4A96209310071) and the National Aeronautics and Space Administration (Grant No. NAGW 3539) for sponsoring the work described in this report. We also wish to acknowledge the technical assistance of undergraduate students, R. Singer, R. Maruyama, and D. Blaschek.

REFERENCES

[1] A. Hasegawa, private communication, 1988.

[2] M. E. Mauel and A. Hasegawa, in *Physics of Space Plasmas (1990), SPI Conference Proceedings and Reprint Series, Number 10*, T. Chang, G. B. Crew, and J. R. Jasperse, eds. (Scientific Publishers, Cambridge, Mass., 1990), p. 461.

[3] M. Schultz, and L. J. Lanzerotti, *Particle Diffusion in the Radiation Belt,* Volume 7, Physics and Chemistry in Space, (Springer-Verlag, New York, 1974).

[4] M. P. Nakada and G. D. Mead, J. Geophys. Res. **70**, 4777 (1965).

[5] T. J. Birmingham, J. Geophys. Res., **74**, 2169 (1969).

[6] T. A. Farley, A. D. Tomassian, and M. Walt, Phys. Rev. Lett., **25**, 47 (1970).

[7] A. D. Tomassian, T. A. Farley, and A. L. Vampola, J. Geophys. Res., **77**, 3441 (1972).

[8] H. P. Warren, A. Bhattacharjee, and M. E. Mauel, Geophys. Res. Lett., **19**, 941 (1992).

[9] A. A. Chan, L. Chen, and R. B. White, Geophys. Res. Lett.,**16**, 1133 (1989).

[10] K. Takahasi, P. Gigbie, and D. Baker, J. Geophys. Res., **90**, 1473 (1985).

[11] X. Li, *et al.*, J. Geophys. Res., **98**, 215 (1993).

[12] L. Chen and A. Hasegawa, J. Geophys. Res., **96**, 1503 (1991).

[13] V. M. Vasyliunas, J. Geophys. Res., **73**, 2839 (1968).

[14] N. Krall, Phys. Fluids, **9**, 820 (1966).

[15] H. Berk, Phys. Fluids, **19**, 1255 (1976).

Observation of Chaotic Particle Transport Induced by Drift-Resonant Fluctuations in a Magnetic Dipole Field

H. P. Warren and M. E. Mauel

Department of Applied Physics, Columbia University, New York, New York 10027

(Received 5 May 1994)

The chaotic radial transport of energetic electrons trapped in a magnetic dipole field has been observed in a laboratory terrella. This transport is driven by multimode, drift-resonant plasma instabilities which are excited by the hot electron population. A transport simulation of energetic electrons interacting with a spectrum of electrostatic waves modeled on the measured fluctuations reproduces temporal features of the experimentally observed radial particle flux.

PACS numbers: 52.25.Fi, 52.35.–g, 94.20.Rr

Charged particles trapped in a dipole magnetic field undergo collisionless radial transport when nonaxisymmetric fluctuations break the third adiabatic invariant ψ, which is proportional to the unperturbed magnetic flux. For example, random variations in the solar wind intensity produce perturbations of the Earth's geomagnetic and convection electric fields which have a broad fluctuation spectrum dominated by low-order azimuthal components [1]. Quasilinear models of the resulting transport have been used to account for the radial profile of radiation belt particles measured with satellites [2,3]. The more general problem of chaotic radial transport driven by nonlinear wave-particle resonances in a dipole magnetic field has been examined by Chan, Chen, and White [4] using Hamiltonian methods. For spectra characterized by multiple discrete modes, the extent of the transport is restricted both in radius and in energy by the existence of specific wave-particle resonances, and the transport need not be quasilinear [5]. Recently, these techniques have been used to study the time evolution of proton phase space distributions in the Earth's magnetosphere induced by a given fluctuation spectrum [6].

In this Letter, we report the first observations of wave-induced chaotic radial transport in a laboratory terrella, the Collisionless Terrella Experiment (CTX). These observations demonstrate a clear relationship between the wave spectrum and the induced resonant transport and provide the first laboratory test of Hamiltonian methods which can used to simulate magnetospheric transport. In addition, we observe wave-particle dynamics in an evolving nonlinear system which is a topic important to transport studies in other magnetically confined plasmas [7].

In CTX, an energetic population of trapped electrons is produced using electron cyclotron resonance heating (ECRH) [8]. The trapped electrons excite quasiperiodic "bursts" of drift-resonant instabilities which we identify as the hot electron interchange mode [9,10]. We find that during these bursts the measured amplitudes, frequencies, and azimuthal mode numbers of the drift-resonant fluctuations meet the conditions required for global chaotic particle transport. During these times, significantly enhanced electron transport is observed with a gridded particle detector and this transport is strongly modulated at the drift frequency of the energetic electrons. At other times, when the instability wave spectrum does not satisfy the conditions for global chaos, no enhanced transport is observed. A numerical simulation of trapped electrons interacting with electrostatic waves modeled on experimental measurements reproduces the modulation depth and frequency of the observed electron flux.

The experiments reported here differ significantly from laboratory terrella previously built to aid the understanding of global magnetospheric structure in the presence of a steady solar wind [11]. The CTX device does not have a "simulated" solar wind, and the low-density, hot electron plasmas produced in CTX are much less collisional than terrella experiments related to the solar wind. The experiments reported here also differ from the "collisionless" terrella experiments conducted in ultrahigh vacuum by Il'in and Il'ina [12]. They demonstrated the breakdown of the first two adiabatic invariants at high energy and observed a reduced adiabatic limit during the application of axisymmetric perturbations. In contrast, we study chaotic particle transport of an energetic belt of trapped electrons induced by low-frequency, nonaxisymmetric waves which preserve the first two adiabatic invariants [13].

The CTX device consists of a dipole magnet suspended mechanically within an aluminum vacuum vessel approximately 1.4 m in diameter. The plasma is created and heated by applying a 1 kW pulse of 2.54 GHz microwaves for approximately 1 sec. Just prior to and during the microwave pulse, hydrogen gas is puffed into the vacuum chamber to control the background neutral pressure. Typically, the initial pressure is ~ 0.3 μTorr and rises during the discharge depending on the duration and number of gas puffs. After a period lasting a few tenths of a second, a ring of hot trapped electrons forms near the flux surface containing the fundamental cyclotron resonance at the equatorial plane of the dipole. This hot electron population is similar to those created with ECRH in other magnetic traps [14]. Figure 1 shows the magnetic field geometry, the approximate location of the energetic electrons, and the representative placement of some of the probe diagnostics discussed below.

0031-9007/95/74(8)/1351(4)$06.00 © 1995 The American Physical Society

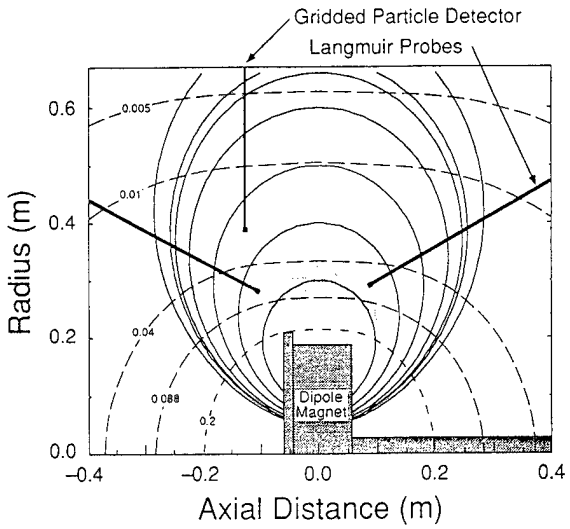

FIG. 1. Magnetic field lines (solid) and $|B|$ contours (dashed) in the CTX device. Heavy lines represent positions of movable probes. The shaded region indicates the approximate location of the energetic electron belt.

The characteristics of the energetic electrons are determined using a krypton proportional counter surrounded by a lead collimator aimed to view the equator of the dipole magnet. Bremsstrahlung emission at energies between 1 and 60 keV is recorded with a histogramming, multichannel analyzer every 50 msec. The energy and intensity of the emission are strongly dependent on the neutral hydrogen pressure. The most intense emission occurs when the pressure is ~ 3 μTorr. The observed distributions are non-Maxwellian, characteristic of microwave-heated electrons [14]. The electrons with energies between 1 and 10 keV are referred to as the "warm" population, and electrons with energies above 10 keV are referred to as the "hot" population. When the microwave power is switched off, the hot population persists for 5–20 msec, defining the discharge "afterglow."

The drift-resonant fluctuations are observed with Langmuir probes and high-impedance floating potential probes situated at five locations within the vacuum vessel. The probes can be repositioned radially to examine the density and potential fluctuations at different flux surfaces. Simultaneous measurements from multiple probes are used to determine the azimuthal mode number and radial mode structure of the fluctuations. We take the amplitude of the floating potential oscillations to be representative of the electrostatic oscillations of the plasma waves. Data from the probes are digitized using transient recorders capable of storing 2^{17} samples at rates of up to 100 MHz. The Langmuir probes are also used to estimate the spatial extent of the hot electron population since the x-ray emission is diminished as the probes contact the resonance region.

The transport of energetic electrons is detected with a movable, gridded particle detector. The detector is a cube with a width of 7 mm and with an angular acceptance of 1.5 sr. The grids are biased to repel ions and electrons with energies less than 100 eV, and the entrance aperture can be rotated with respect to the local magnetic field vector. For the measurements described here, the entrance was oriented perpendicular to the magnetic field excluding electrons with energies below approximately 3 keV due to gyroradius effects. Figure 1 shows the typical location of the gridded particle detector, slightly offset from the dipole's equatorial plane.

When an intense hot electron population is produced, drift-resonant fluctuations ($\omega \sim \omega_{dh}$) are observed both while the ECR heating is on and in the afterglow. During heating, the fluctuations occur in quasiperiodic bursts lasting approximately 300–500 μsec and having initial growth rate of $\sim 1/50$ μsec. The rate at which the bursts appear is a function of the hot electron intensity: at the highest observed energies the bursts become almost continuous, while at lower energies the interval between bursts increases. During the afterglow, the drift-resonant oscillations have a slower growth rate, $\sim 1/100$ μsec, and persist for several milliseconds. At both times, the observed instabilities propagate azimuthally in the direction of the electron ∇B drift, are flutelike with a constant phase along a field line, and have a broad radial structure extending throughout the plasma. The saturated amplitudes of the floating potential oscillations present at both times are also similar, typically 100–200 V, with the larger amplitudes resulting from the most energetic electron populations.

The time evolution of the spectral content of the waves was examined by computing the spectrogram of the potential fluctuations. Figure 2 shows the spectrogram of a floating potential probe signal taken at the end of the heating phase and extending into the afterglow. The quasiperiodic bursts consist of multiple frequencies typically below $f \leq 2$ MHz although some particularly intense bursts have reached frequencies as high as 5 MHz. These frequencies drift resonate with the warm electrons. The spectrogram also shows that the wave frequencies generally increase in time or "chirp." The rising tones are most evident during the intense coherent modes seen during the afterglow when the frequencies range from 1 to 10 MHz, resonating with the hot electrons. The chirp rate during the afterglow is sensitive to the intensity of the hot electrons: at high intensity the chirp rate is fast, and at lower intensities the rise in frequency is slower.

The Fourier transform of the correlation between Langmuir probes located at the same radius but separated in azimuth was computed in order to determine the azimuthal mode number of the various frequencies. The azimuthal mode number of the quasiperiodic bursts is usually limited to $m = 1$ except in the most intense bursts where modes with $m = 2$ are observed at high frequencies. During the afterglow $m \leq 6$ and several waves are observed simultaneously having the same m number.

We believe the drift-resonant fluctuations observed in CTX are related to the hot electron interchange

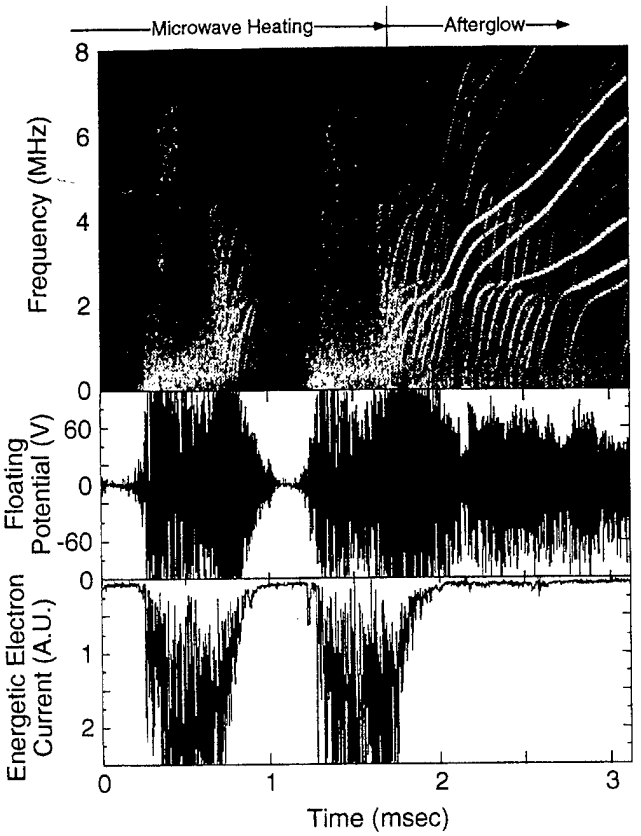

FIG. 2. Simultaneous measurements of drift-resonant fluctuations and energetic electrons near the end of the microwave heating pulse. Spectrogram (top) illustrates the multimode fluctuations with rising tones. Floating potential oscillations (middle) illustrate intense, quasiperiodic bursts correlated with rapid increase of electron flux (bottom) to the gridded particle detector.

instability (HEI) that has been observed in ECRH heated magnetic mirror experiments [8,15]. The linear dispersion relation for low-frequency ($\omega \ll \omega_{ci}$) electrostatic flute modes was developed by Krall [9] and extended to high-frequency modes ($\omega \sim \omega_{ci}$) by Berk [10]. These theories, which assume a monoenergetic distribution of hot electrons and slab geometry, predict instability when the energetic electron fraction exceeds some critical value.

When the linear theory is extended to include a distribution of velocities and dipole geometry, it predicts unstable modes with real frequency $\omega < m\bar{\omega}_{dh}/2$, where $\bar{\omega}_{dh}$ is the precessional drift frequency averaged over the distribution of energies. The fact that the real frequency of the instability is below the average drift frequency of the energetic electron population suggests a possible explanation of the chirping observed in the experiment. The instability drives transport in the lower energy particles first. The remaining distribution has a higher average energy which leads to a higher real frequency for the instability.

Coincident with the bursts of wave activity during ECRH, we observe increases in the flux of energetic electrons to the gridded particle detector indicating rapid radial transport. As shown in Fig. 2, the electron current rises significantly above the background level during the bursts. No transport is associated with the high-frequency wave activity observed in the afterglow. At this time, the measured wave spectrum produces isolated bands of stochastic radial motion not capable of global transport.

A feature of particular interest in Fig. 2 is the persistent modulation of the electron flux near the precessional drift frequency of the hot electrons. This type of behavior is suggestive of "drift echoes" which have been observed in satellite measurements of energetic particle transport [16]. Drift echoes result from the sudden injection of energetic particles at a narrow range of longitude. If the injected particles differ from the background population, each time they drift past the satellite a change in the particle flux is observed.

We conjecture that a similar process is responsible for the modulation of the observed particle flux in CTX. As we will show, the experimentally measured wave spectra are consistent with the observation of chaotic particle motion. The decorrelation time, however, is long with respect to a drift period. Thus inhomogeneities in phase space propagate azimuthally and lead to temporal variations in the particle flux.

The interaction of energetic electrons with the electrostatic waves observed in the experiment can be described by the guiding center drift Hamiltonian [17]

$$H = \tfrac{1}{2} m_e \rho_{\parallel}^2 B^2 + \mu B - e\Phi , \qquad (1)$$

where $\rho_{\parallel} \equiv v_{\parallel}/B$, and $\mu \equiv m_e v_{\perp}^2/2B$ is the first adiabatic invariant, the magnetic moment. The canonical variables, $(\rho_{\parallel}, \chi)$ and (ψ, φ), of the Hamiltonian are defined by $\mathbf{B} = \nabla\psi \times \nabla\varphi = \nabla\chi$, where $\psi = M\sin^2\theta/r$, $\chi = M\cos\theta/r^2$, $M \equiv B_0 R_0^3$ is the moment of the dipole magnet, and (r, φ, θ) are spherical coordinates. For simplicity, we examine only equatorial particles with $\rho_{\parallel} = \chi = 0$.

The fluctuations present during the quasiperiodic bursts are modeled as a sum of $m = 1$ fixed-frequency traveling waves,

$$\Phi(\varphi, t) = \sum_{l=1}^{N} \Phi_l \cos(\varphi - \omega_l t + \varphi_l) . \qquad (2)$$

The frequencies and relative amplitudes of the waves are taken directly from experimental observations. Each wave in the series forms a drift island of width $\Delta\psi_l \simeq (2e\Phi_l \psi_l^r/\omega_l)^{1/2}$, where ψ_l^r is the resonant surface defined by $\omega_d(\psi_l^r) = \omega_l$. When multiple drift-resonant waves have sufficiently large amplitudes, the islands will overlap and lead to chaotic transport [4]. This condition for chaos is visualized by constructing a Poincaré surface of section by plotting particle coordinates, (ψ, φ), at multiples of the mapping time defined by the least common multiple of the periods of the wave motion [18].

For a series of waves modeled on the initial part of an instability burst, drift resonances exist for electrons with energies between 3 and 10 keV (i.e., the warm electrons) from the center of the hot electron ring to the wall of

the vacuum chamber. Examination of the phase space portraits indicates that when the wave amplitude is above 75 V there are no encircling Kolmogorov-Arnol'd-Moser (KAM) surfaces preventing global transport. This implies that for the wave spectra measured in the experiment the conditions for chaotic particle transport are met.

In order to demonstrate the relationship between the modulation of the current and the spectral content of the observed fluctuations, we have simulated the time evolution of the flux of electrons to a small region of phase space which represents the particle detector. In the simulation, we randomly selected an ensemble of 5000 particle trajectories that are at the "gridded particle detector" at time $t = T$. Using spectral information from the experiment to construct the electrostatic potential in the form of Eq. (2), the equations of motion are integrated backwards in time from $t = T$ to $t = 0$, and the probability that the trajectory came from an assumed hot electron distribution, $F_h(\mu, \psi)$, is computed. This process is repeated in order to compute the current as a function of time.

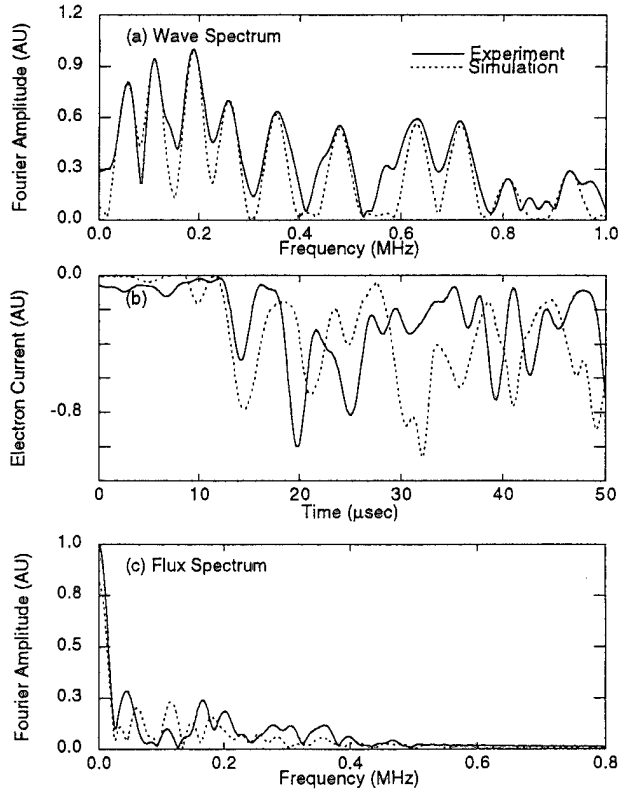

FIG. 3. A comparison of experiment and a simulation which reproduces the gross frequency and depth of the modulation. (a) Frequency spectrum of the electrostatic waves, (b) detected energetic electron flux, and (c) Fourier transform of the detected flux.

The result of one such simulation is presented in Fig. 3. Here the frequency spectrum, the energetic electron flux to the detector, and the Fourier transform of the flux are shown for both the experiment and the simulation. The simulation reproduces the gross frequency and depth of the modulation present in the observed electron flux. However, since the phases of the measured waves and other profile parameters are unknown, the simulation cannot be expected to reproduce the particle flux exactly.

In summary, radial transport driven by drift-resonant fluctuations has been observed in a laboratory terrella experiment when the spectral characteristics of the observed fluctuations meet the conditions for chaotic particle motion. Strong modulation in the observed electron flux is reproduced in a numerical transport simulation.

We would like to acknowledge the advice of Akira Hasegawa during the design of CTX and recognize the contributions of R. Singer and R. Maruyama. This work was supported by AFOSR Grant No. F4A96209310071, NASA Grant No. NAGW-3539, and NSF Grant No. ATM-91–11396.

[1] M. Shulz and L. J. Lanzerotti, *Particle Diffusion in the Radiation Belts* (Springer-Verlag, New York, 1974).
[2] M. P. Nakada and G. D. Mead, J. Geophys. Res. **70**, 4777 (1965).
[3] T. A. Farley, A. D. Tomassain, and M. Walt, Phys. Rev. Lett. **25**, 47 (1970).
[4] A. Chan, L. Chen, and R. White, Geophys. Res. Lett. **16**, 1133 (1989).
[5] A. B. Rechester and R. B. White, Phys. Rev. Lett. **44**, 1586 (1980).
[6] M. W. Chen, L. R. Lyons, and M. Schulz, J. Geophys. Res. **99**, 5745 (1994).
[7] C. T. Hsu *et al.*, Phys. Rev. Lett. **72**, 2503 (1994).
[8] M. Mauel, H. Wârren, and A. Hasegawa, IEEE Trans. Plasma Sci. **20**, 626 (1992).
[9] N. A. Krall, Phys. Fluids **9**, 820 (1966).
[10] H. L. Berk, Phys. Fluids **19**, 1255 (1976).
[11] K. Schindler, Rev. Geophys. **7**, 51 (1969).
[12] V. D. Il'in and A. N. Il'ina, Sov. Phys. JETP **43**, 661 (1976); **45**, 514 (1977).
[13] H. P. Warren, A. Bhattacharjee, and M. E. Mauel, Geophys. Res. Lett. **19**, 941 (1992).
[14] For example, R. A. Dandl *et al.*, Nucl. Fusion **4**, 344 (1964), or H. Ikegami *et al.*, Phys. Rev. Lett. **19**, 778 (1967).
[15] S. Hiroe *et al.*, Phys. Fluids **27**, 1019 (1984).
[16] L. J. Lanzerotti, C. S. Roberts, and W. L. Brown, J. Geophys. Res. **72**, 5893 (1967).
[17] A. H. Boozer, Phys. Fluids **9**, 904 (1980).
[18] A. G. Kornienko *et al.*, Phys. Lett. A **158**, 398 (1991).

Wave-induced chaotic radial transport of energetic electrons in a laboratory terrella experiment

H. P. Warren[a] and M. E. Mauel
Department of Applied Physics, Columbia University, New York, New York 10027

(Received 14 June 1995; accepted 11 August 1995)

This paper reports the observation of wave-induced chaotic radial transport of energetic electrons in a laboratory terrella, the Collisionless Terrella Experiment (CTX) [H. P. Warren and M. E. Mauel, Phys. Rev. Lett. **74**, 1351 (1995)]. Electron cyclotron resonance heating is used to create a localized population of energetic electrons which excite the hot electron interchange instability. The electrostatic fluctuations driven by this instability have time-evolving spectra which resonate with the precessional drift motion of the hot electrons. We have established that the amplitude, frequency, and azimuthal mode number of the observed instabilities meet the conditions for the onset of chaotic particle motion. Electron transport is observed with a gridded particle detector. Increases in the flux of energetic particles to the detector are well correlated with the presence of fluctuations which meet the conditions for global chaos. Greatly diminished transport is observed when the fluctuations lead to thin, radially localized bands of chaos. The flux of energetic electrons to the detector is strongly modulated. By examining time-dependent Hamiltonian phase space flows, the modulation is shown to be the result of phase space correlations. A transport simulation based on the Hamiltonian motion of energetic electrons reproduces the frequency and modulation depth of the observed electron flux and allows for comparison between the Hamiltonian and quasilinear descriptions of transport. © *1995 American Institute of Physics.*

I. INTRODUCTION

Chaotic radial transport plays a central role in the formation and evolution of energetic particle populations trapped in planetary magnetospheres. Of primary importance to models of the Earth's magnetosphere is transport driven by random variations in the solar wind intensity and the interplanetary magnetic field which couple to the Earth's geomagnetic and convection electric fields.[1] Sudden, large-scale variations which have a decay time on the order of a precessional drift period can lead to resonant wave-particle interactions and the breaking of the third adiabatic invariant, ψ. The resulting transport is an essential mechanism for both the injection of particles into the magnetosphere and their acceleration.[2]

The dynamics of energetic particle populations in the Earth's magnetosphere may be of practical importance. For example, sudden increases in relativistic electron fluxes associated with geomagnetic storms are thought to be responsible for the failure of electronic components on spacecraft.[3,4] It has also been suggested that the precipitation of relativistic electrons can affect levels of stratospheric ozone by creating hydrogen and nitrogen oxides which catalytically destroy ozone.[5]

Quasilinear models have been used to account for the radial profile of radiation belt particles measured with satellites[6,7] as well as the stormtime enhancement and quiet time evolution of ring current particles.[8–10] Quasilinear theory is, however, only an approximate statistical description of many transport processes. Phase space structures often persist even in the presence of symmetry-breaking perturbations and the dispersion predicted by quasilinear theory is not observed.[11] Even when phase space is almost completely stochastic, deviations from quasilinear theory are possible.[12] This is especially relevant to many physically realized situations where the conditions necessary for the validity of quasilinear theory are not strictly met.

A more complete description of particle transport is given by the guiding center drift Hamiltonian.[13,14] The chaotic radial transport of ring current protons driven by magnetospheric hydrodynamic waves has been examined by Chan, Chen, and White[15] using Hamiltonian methods. For spectra characterized by multiple discrete modes, the extent of the transport is restricted both in radius and in energy by the existence of specific wave-particle resonances. A Hamiltonian approach has also been used to model the enhancement of ring current protons due to impulsive stormtime variations in the Earth's convection electric field by Chen *et al.*[16–18]

In an earlier paper,[19] we reported the first observation of wave-induced chaotic radial transport in a laboratory terrella experiment, the Collisionless Terrella Experiment (CTX). One of the primary goals of CTX is to study the process of chaotic radial transport in dipole magnetic fields. In particular, these experiments investigate the relationship between fluctuation spectra and models of energetic particle transport and provide the first laboratory tests of Hamiltonian methods which can be used to simulate magnetospheric transport. This paper presents a more complete description of the experimental observations and reports the detailed results of numerical simulations which model transport processes observed in the experiment.

In CTX, an energetic population of trapped electrons is produced using electron cyclotron resonance heating (ECRH).[20] The trapped electrons excite quasiperiodic "bursts" of drift-resonant electrostatic fluctuations which we identify as the hot electron interchange instability.[21,22] Al-

[a]Present address: Naval Research Laboratory, Washington, D.C. 20375.

though the hot electron interchange instability is not excited naturally in planetary magnetospheres, the process by which nonaxisymmetric fluctuations induce chaotic radial transport of collisionless, trapped particles is independent of wave dispersion characteristics. Instead, chaotic radial transport occurs only when the spectral and spatial characteristics of the fluctuations meet specific requirements, and these requirements are documented in the laboratory using the hot electron interchange mode. We note further that the drift-resonant fluctuations drive strongly chaotic radial motion which preserves the first two adiabatic invariants μ and J.[23] Thus the experimentally observed transport processes are relevant to magnetically trapped magnetospheric particles drift-resonant with fluctuations which have low azimuthal mode numbers.

By computing Poincaré surfaces of section we find that during the instability bursts the measured amplitudes, frequencies, and azimuthal mode numbers of the drift-resonant fluctuations meet the conditions required for global chaotic particle transport. During the instability bursts significantly enhanced electron transport is observed with a gridded particle detector. The observed transport is strongly modulated at frequencies related to the precessional drift-frequency of the energetic electrons. At other times, when the instability wave spectrum does not satisfy the conditions for global chaos, no enhanced transport is observed.

The Poincaré surfaces of section which model particle motion during the instability bursts show that phase space is strongly chaotic and indicate that quasilinear theory should be applicable. However, in order to model the spatial and temporal evolution of particle fluxes observed in the experiment, computer simulations using the guiding center drift Hamiltonian are necessary. The first of these simulations involves visualizing the Hamiltonian flow of phase space as a function of time so that the evolution of equatorial particle trajectories can be examined. By examining animations of phase space flows, the observed modulation in the electron flux can be understood in terms of phase-space correlations. The second simulation involves computing the flux of equatorial particle trajectories to a small region of phase space which represents the particle detector. This simulation reproduces the modulation depth and frequency of the experimentally observed electron flux. When the results of the Hamiltonian transport simulation are compared with a simple quasilinear transport model we find that the predictions of the quasilinear model do not reproduce the experimentally observed modulations and underestimate the maximum particle fluxes measured in the experiment.

An interesting aspect of the experimental measurements we report here is the complicated evolution of the frequency content of the observed fluctuations. Both at the end of a quasiperiodic burst and during the afterglow, when the conditions for global chaos are not met, the wave-frequency increases in time or "chirps." Such behavior is often exhibited by evolving nonlinear systems and is a topic important both to nonlinear dynamics in general and to specific transport studies in other magnetically confined plasmas.[24]

The Collisionless Terrella Experiment differs significantly from previous terrella experiments designed to study the interaction of a steady solar wind with the geomagnetic field.[25] For these experiments, the CTX device employed only a pure dipole magnetic field. CTX does not have a "simulated" solar wind and magnetic field coils that can be used to create a quasimagnetopause were not used. The plasmas produced in CTX are low-density, hot electron plasmas which do not have the same magnetohydrodynamic properties as those used to study the solar wind. The experiments reported here also differ from the "collisionless" terrella experiments conducted in ultra-high vacuum by Il'in and Il'ina.[26,27] They demonstrated the breakdown of the first two adiabatic invariants at high energy and observed a reduced adiabatic limit during the application of axisymmetric perturbations.

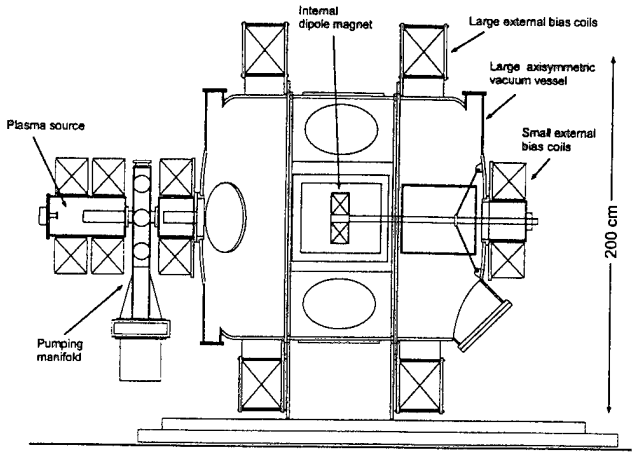

FIG. 1. A schematic diagram of the Collisionless Terrella Experiment showing the vacuum vessel and the dipole magnet.

This paper is organized in the following manner. A description of the experimental device and the plasma diagnostics is given in Sec. II. The results of the experiments related to the observation of hot electron plasmas, drift-resonant fluctuations, and chaotic radial transport are given in Sec. III. The results of computer simulations which model the resulting phase-space topology using Poincaré surfaces of section, the evolution of Hamiltonian phase space flows, and the flux of energetic electrons to the particle detector are given in Sec. IV. Finally, a summary and discussion of the results are given in Sec. V.

II. THE COLLISIONLESS TERRELLA EXPERIMENT

The CTX experimental device consists of a dipole electromagnet suspended mechanically in an axisymmetric aluminum vacuum vessel 1.9 cm thick and 137 cm in diameter. As shown in Fig. 1, the magnet is supported by a stainless steel casing which houses the electrical and cooling leads. The strength of the dipole magnetic field is approximately 15 kG at the face of the magnet and falls off to less than 50 G at the wall. The magnetic field lines and mod-B surfaces of the dipole magnet used in CTX and are shown in Fig. 2.

Plasmas are created by breaking down and heating hydrogen gas using microwaves which resonate with the electron gyromotion (this is electron cyclotron resonance heating or ECRH). The microwave source is a continuous wave mag-

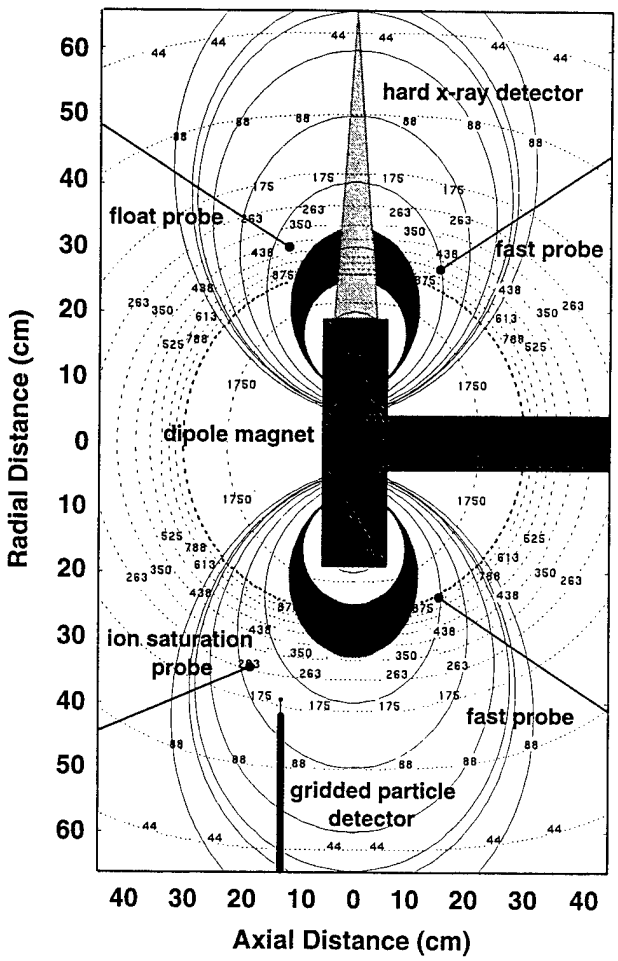

FIG. 2. The magnetic field geometry of CTX. The solid lines represent magnetic field lines and the dotted lines represent surfaces of constant magnetic field. The location of several plasma diagnostics and the approximate location of the hot electron ring are also shown.

netron which has a peak output of 1500 W at a frequency of 2.45GHz. The rate of heating is proportional to the amount of time that an individual electron is resonant with a wave and the heating will be strongest for particles mirroring along the surface defined by $B_0 \equiv B \simeq 875$ G. The equatorial field strength of the dipole magnet has this value at a radial distance of $R_0 \equiv R \simeq 27$ cm and this defines the center of the hot electron "ring" or "artificial radiation belt."

Plasma diagnostics include a series of Langmuir and high impedance floating potential probes situated at five locations throughout the vacuum vessel. The probes can be repositioned radially to examine the density and potential fluctuations at different flux surfaces. Multiple probe measurements allow the direction of propagation, the azimuthal mode number and radial mode structure of the observed fluctuations to be determined. Data from the electrostatic probes are digitized using 8-bit, high-speed transient recorders which are capable of storing 2^{17} samples at rates of up to 100 million samples per second.

A krypton proportional counter is positioned at the midplane of the magnetic field and is used to measure hard x-ray emission at energies between 1–60 keV. The counter is collimated to view the equatorial region of the dipole magnetic field. X-ray spectra are collected and stored at 50 ms intervals during the plasma discharge using a multichannel analyzer and a histogramming memory unit.

A movable gridded particle detector is located inside the vacuum chamber 13 cm above the equatorial midplane. The particle detector is supported by a metal rod which extends from the vacuum chamber wall and can be repositioned to measure particle flux at different radial positions. The detector consists of a small square metal box 1 cm on a side with a 0.23 cm^2 opening centered on one face. The box contains a series of grids which are biased to repel ions and electrons with energies less than 100 eV.

The entrance aperture of the probe can be rotated with respect to the local magnetic field vector. For the measurements described here, the entrance was oriented perpendicular to the magnetic field. Gryoradius effects determine the minimum particle energy detected by the probe. The minimum detectable energy ranges from 700 eV to 2.3 keV depending on the radial position of the probe. Since the probe is located above the midplane of the magnetic field, only particles with a sufficiently small equatorial pitch angle will be detected by the particle detector. The maximum pitch angle ranges from 54° to 65°.

III. EXPERIMENTAL OBSERVATIONS

The operation of the Collisionless Terrella Experiment is relatively simple. After the dipole electromagnet is turned on, hydrogen gas is puffed into the vacuum chamber and heated with microwaves. After a specified time the microwave power is turned off and the hot electron "afterglow" of the plasma discharge can be examined. Computer controlled timing signals define intervals when the high speed transient recorders store diagnostic signals.

A. Characterization of hot electron plasmas

The intensity of the hot electron population is characterized by the hard x-ray emission produced by electron-ion and electron-neutral bremsstralung. A fraction of these x-rays are detected with the krypton proportional counter. The measured x-ray spectra indicate the formation of a magnetically trapped, energetic electron population with energies between 1–60 keV. The observed distributions are non-Maxwellian and more closely resemble power-law distributions often associated with wave-heated space plasmas.[28] There are two distinct subpopulations within the electron distribution. There is a "warm" population with energies between 1–10 keV, and a "hot" population, which is more sensitive to variations in background pressure, with energies above 10 keV.

As shown in Fig. 3, the distribution of energetic electrons is dependent on the hydrogen fill pressure. Whenever the peak pressure falls either above or below the optimal value the hot electron population is reduced. A similar dependence of hot electron intensity on background pressure has been reported in other ECR heated magnetic mirror experiments.[29]

The spatial extent of the energetic electron population is established by using a Langmuir probe as a "limiter" or "skimmer" and moving it into the resonance region until the

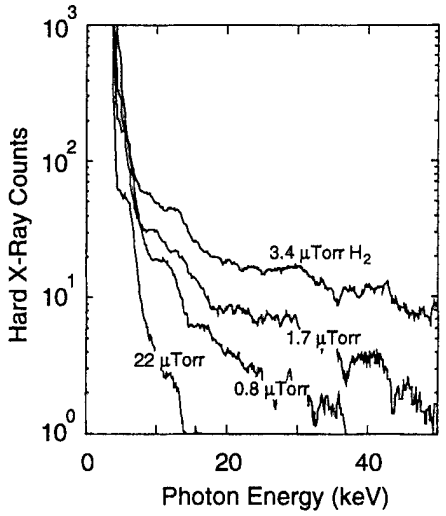

FIG. 3. Hard x-ray emission as a function of hydrogen fill pressure. If the peak pressure falls above or below the optimal value, the hot electron energy will be reduced.

x-ray emission is significantly diminished. Plasmas without a significant hot electron population can also be created by decreasing the dipole's magnetic field. When the field strength is sufficiently small, the resonant surface coincides with the surface of the magnet and no hard x-ray emission is observed.

B. Observation of drift-resonant fluctuations

When a sufficiently intense hot electron population is present, electrostatic fluctuations which resonate with the precessional drift-motion of the hot electrons ($\omega \sim \omega_{dh}$) are detected by the Langmuir probes. Fluctuations are observed both while the ECR heating is on and in the afterglow of the plasma discharge, when the heating has been turned off.

While the ECR heating is on, quasiperiodic bursts of drift-resonant fluctuations with a growth rate of ~50 μs and which last approximately 300–700 μs are observed. The rate at which the instability bursts appear is a strong function of the hot electron energy: at the highest observed energies the bursts become almost continuous, at lower energies the time between bursts increases to intervals as long as several milliseconds. During the afterglow, the observed fluctuations have a slower growth rate, ~100 μs, and persist for several milliseconds.

During both the ECRH and in the afterglow, the observed instabilities propagate azimuthally in the direction of the curvature and ∇B drift motions of the electrons. The observed instabilities are flute modes with no structure along a field line and have only a weak radial dependence. The saturated wave amplitudes observed during both the ECRH and the afterglow are also similar, typically 100–200 V, with the larger amplitudes resulting from the most energetic electron populations.

The time evolution of the frequency content of the observed fluctuations can be examined by computing the time-frequency domain (TFD) of the Langmuir probe signals. One of the simplest and most intuitive algorithms for computing the TFD is the short-time Fourier transform or

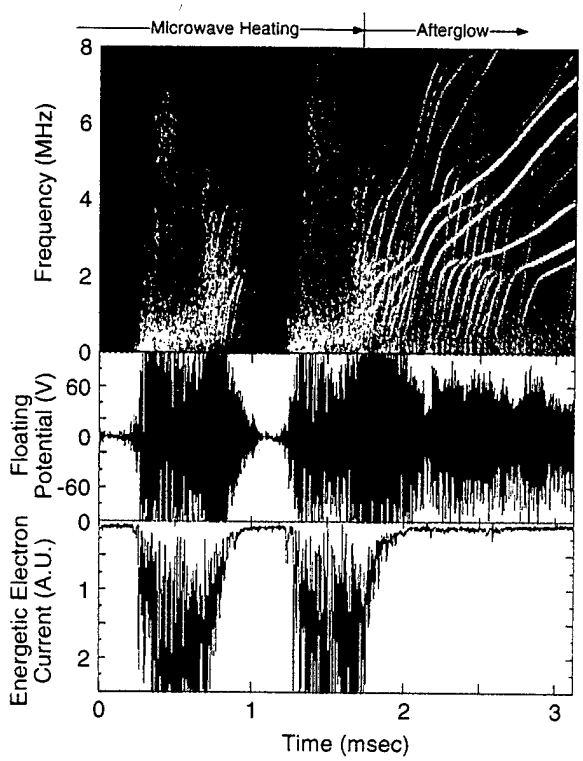

FIG. 4. The flux of energetic electrons to the gridded particle detector, the measured floating potential, and a color contour plot of the corresponding spectrogram. The observation of transport with the particle detector is correlated with the presence of fluctuations which satisfy the conditions for global chaos. The measurements begin during the ECRH phase and extend into the afterglow of the plasma discharge.

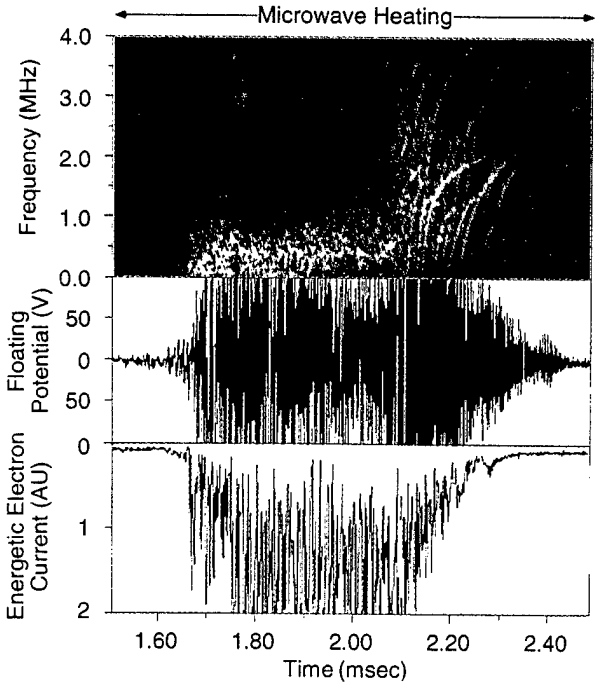

FIG. 5. A detailed spectrogram of the first instability burst shown in 4. For the first 450 μs of the burst the fluctuations are incoherent and resonate with the "warm" part of the energetic particle population and enhanced transport is observed. For the final 200 μs the fluctuations are more coherent, the instability "chirps" upward in frequency, and the transport is diminished.

spectrogram.[30] An example of a discharge in which drift-resonant fluctuations have been observed is shown in Fig. 4 where a floating potential probe signal and its spectrogram are displayed. The measurement was taken at the end of the heating phase and extends into the afterglow. This Fourier analysis indicates that the quasiperiodic bursts consist of incoherent, broad-band fluctuations with frequencies typically below 2 MHz, although some particularly intense bursts have frequencies as high as 5 MHz. The instabilities observed during the afterglow are more coherent than those observed during the ECRH and have a higher range of frequencies, $f\sim 1-12$ MHz.

Inspection of the spectrogram shown in Fig. 4 reveals that the wave frequency evolves in a very complicated way both during the ECRH and in the afterglow. Fig. 5 shows the spectrogram of a single burst in detail and it is clearly seen that the quasiperiodic fluctuations consist of two different stages. During the initial stage of the burst, lasting 300 μsec, the fluctuations have frequencies below 1 MHz and appear to be incoherent. At the end of the burst, the fluctuations become more coherent and rapidly increase in frequency or "chirp." During the afterglow only coherent, rising tone fluctuations are observed.

The fluctuations observed during the ECRH also differ from those observed in the afterglow in terms of their azimuthal mode number. During the quasiperiodic instability bursts, the azimuthal mode number is usually limited to $m=1$, except in the most intense bursts when some high-frequency, $m=2$ modes are observed. In the afterglow, $m\leqslant 6$, and there are often multiple waves with the same azimuthal mode number present simultaneously.

The drift-resonant fluctuations observed in the experiment are related to the hot electron interchange instability that has been observed in other ECR heated magnetic mirror experiments.[31,32] The linear theory of the hot electron interchange, assuming slab geometry and a monoenergetic distribution of electrons, was developed by Krall for $\omega\ll\omega_{ci}$[21] and by Berk[22] for $\omega\sim\omega_{ci}$ and the observed fluctuations are generally similar to those predicted by these simple theories.

C. Observation of energetic electron transport

Coincident with the bursts of wave activity observed during the ECRH are increases in the flux of energetic electrons to the gridded particle detector. The characteristics of the transport measurements suggest a definite relationship between the spectral content of the waves and the observed energetic particle flux. As shown in both Fig. 4 and Fig. 5, the electron current rises significantly above the background level during the low-frequency, incoherent phase of the quasiperiodic bursts. Transport is diminished during the coherent phase of a burst when the frequency of the instability "chirps" upward. No enhanced transport is associated with the high-frequency, coherent modes observed during the afterglow. Also note that the observed energetic electron flux is strongly modulated at frequencies related to the precessional drift-frequency of the "warm" electrons. The connection between the fluctuation spectrum and energetic electron transport will be examined in detail in the next section.

IV. WAVE-INDUCED PARTICLE TRANSPORT

The interaction of energetic electrons with the drift-resonant electrostatic waves observed in the experiment can be described by the guiding center drift Hamiltonian[14]:

$$\mathscr{H}=\frac{m_e c}{2e}\rho_{\|}^2 B^2+\mu\frac{cB}{e}-c\Phi, \tag{1}$$

where m_e and e are the electron mass and charge, c is the speed of light, B is the magnitude of the dipole magnetic field, $\mu\equiv mv_{\perp}^2/2B$ is the magnetic moment, $\rho_{\|}\equiv v_{\|}/B$, and Φ is the electrostatic potential. For a curl-free magnetic field, the canonical coordinates of the guiding center drift Hamiltonian, $(\rho_{\|},\chi)$, and (ψ,φ), are essentially the magnetic coordinates defined by: $\mathbf{B}=\nabla\psi\times\nabla\varphi=\nabla\chi$. The function $\psi=M\sin^2\theta/r$ is proportional to the magnetic flux bounded by a field line and the function $\chi=M\cos\theta/r^2$ is related to the distance along a field line. Note that $M\equiv B_0R_0^3$ is the moment of the dipole magnet and (r,φ,θ) are spherical coordinates. Also note that the guiding center drift Hamiltonian is appropriate since energetic electrons produced in the experiment are nonrelativistic.

When the particle motion is confined to the equatorial midplane of the magnetic field, $\rho_{\|}=\chi=0$, $B=B(\psi)$, and the equations of motion are reduced to a particularly simple form:

$$\begin{aligned}\frac{d\psi}{dt}&=-\frac{\partial\mathscr{H}}{\partial\varphi}=c\frac{\partial\Phi}{\partial\varphi},\\ \frac{d\varphi}{dt}&=\frac{\partial\mathscr{H}}{\partial\psi}=\mu\frac{c}{e}\frac{\partial B}{\partial\psi}-c\frac{\partial\Phi}{\partial\psi}.\end{aligned} \tag{2}$$

These are the equations that will be considered in this paper. For equatorial particles the precessional drift frequency is:

$$\omega_{dh}=\frac{3c\mu B}{e\psi}=\frac{3\mu B_0}{m_e\Omega_0R^2}, \tag{3}$$

where $\Omega_0=eB_0/m_ec$ and R is the particle's radial position. In general, particles will have a finite parallel velocity and will not be confined to the equatorial midplane. However, the precessional drift-frequency is only weakly dependent on pitch angle[2] and the results presented here would not be significantly altered by including parallel velocity.

The experimental observations described in the previous section indicate that the observed fluctuations can be modeled as a sum of traveling waves of the form:

$$\Phi(\varphi,t)=\frac{\Phi_0}{\sqrt{\mathscr{S}}}\sum_{m,l}^{N}a_{ml}\cos(m\varphi-\omega_l t+\varphi_l), \tag{4}$$

where $\mathscr{S}=\Sigma_{ml}|a_{ml}|^2$. The relative amplitudes, azimuthal mode numbers, frequencies, and phases are determined from the Fourier analysis of experimentally measured Langmuir probe signals.

For a single electrostatic wave the island half-width, $\Delta\psi$, at the resonant surface defined by $\omega_l-m\omega_{dh}(\mu,\psi_{lm}^r)=0$, is given by the solution to the equation:

$$(\Delta\psi)^2\left[1+\frac{\Delta\psi}{3\psi_{lm}^r}\right]-\frac{2c\Phi_0\psi_{lm}^r}{\omega_{dh}}=0. \tag{5}$$

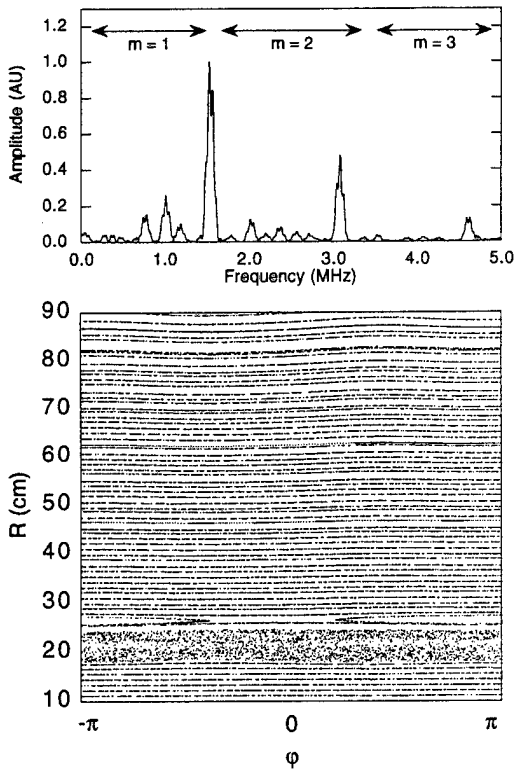

FIG. 6. A spectrum of waves taken from experimental measurements of fluctuations observed during the afterglow and a corresponding Poincaré surface of section. The chaos is localized to a thin, radially localized band near the hot electron ring and consistent with the absence of electron flux to the particle detector. The magnetic moment, μ, is chosen so that $\mu B_0 = 10$ keV; other parameters are $\Phi_0 = 150$ V, $N = 10$, and $T_M = 100$ μs. Note that the surface of section is plotted in (R, φ) coordinates.

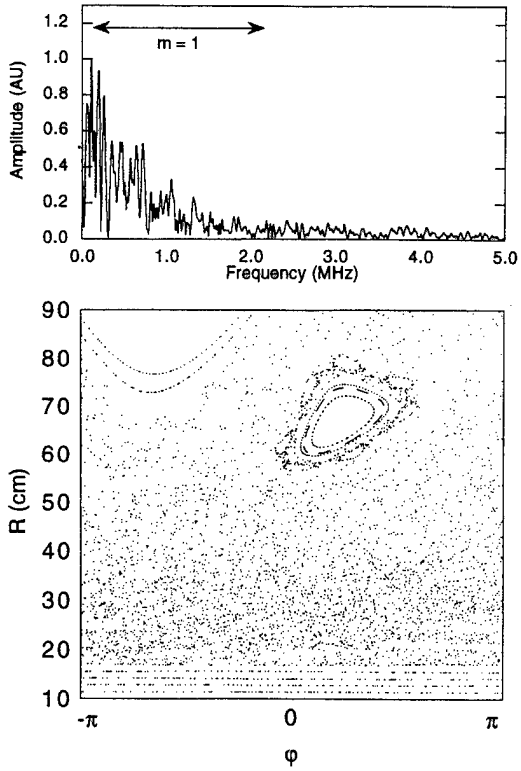

FIG. 7. A spectrum of waves taken from experimental measurements of fluctuations observed during the ECRH and a corresponding Poincaré surface of section. The fluctuations spectrum leads to chaos in ψ over the radial extent of the plasma. The magnetic moment, μ, is chosen so that $\mu B_0 = 4$ keV; other parameters are $\Phi_0 = 150$ V, $N = 10$, and $T_M = 100$ μs. Note that the surface of section is plotted in (R, φ) coordinates.

When the islands are small they are approximately symmetric and the island half-width is given by $\Delta\psi \simeq (2c\Phi_0\psi^r_{lm}/\omega_{dh})^{1/2}$. When multiple drift-resonant waves with sufficiently large amplitude are present, islands will overlap and lead to chaotic transport.

The topological properties of particle trajectories can be determined by constructing the Poincaré surface of section which is formed by plotting (ψ, φ) at multiples of the mapping time. For a spectrum of waves, the mapping time is the least common multiple of the periods of the wave motion: $T_M = \mathrm{LCM}(T_l = 2\pi/\omega_l)$.[33] In practice, the frequencies are rounded off to keep T_M from becoming too large. The equations of motion are integrated using the Gragg-Burlish-Stoer implementation of the Richardson extrapolation.[34]

A. Chaos during the afterglow

An example of a fluctuation spectrum used to model the electrostatic potential observed during the afterglow is shown in Fig. 6a. For electrons with energies above $\mu B_0 \approx 10$ keV this spectrum leads to resonant wave-particle interactions. However, as the Poincaré surface of section shown in Fig. 6b illustrates, the thin bands of chaos in ψ are limited to the proximity of the hot electron ring and do not extend out to the radial location of the gridded particle detector unless the amplitude of the potential fluctuations is very large.

B. Chaos during the quasiperiodic bursts

A number of model fluctuation spectra have been constructed to model the potential fluctuations observed during the ECRH. For the wave spectrum shown in Fig. 7a, which is taken from the initial part of an instability burst, drift-resonances exist for electrons with energies between 1–10 keV (i.e., the "warm" electrons) from the center of the hot electron ring to the wall of the vacuum chamber. As shown in Fig. 7b, this spectrum of waves leads to chaos in ψ over the radial extent of the plasma. Examination of the phase space portraits indicates that when the wave amplitude is above ≈75 V, there are no encircling Kolmogorov-Arnol'd-Moser (KAM) surfaces[35] preventing global transport.

As shown in Fig. 5, the quasiperiodic bursts observed during the ECRH evolve into more coherent modes that rapidly chirp upward to higher frequencies. By the end of a burst these modes have begun to resemble modes observed in the afterglow. Coincident with the transition from incoherence to coherence, the observed particle flux to the detector begins to decline. This is understandable in light of the Poincaré surfaces of section shown in Fig. 6 and Fig. 7. The higher frequency waves lead to chaos that is confined to narrow regions of phase space. This observation also suggests that there is a complicated relationship between the frequency content of the instability, particle transport, and the coherence of the fluctuations.

C. Hamiltonian phase space flows

The transport associated with the quasiperiodic bursts observed during the ECRH is strongly modulated. This type of behavior is suggestive of "drift echoes" which have been observed in satellite measurements of energetic particle transport.[36,37] Drift echoes result from the sudden injection of energetic particles at a narrow range of longitude. If the injected particles differ from the background population, each time they drift past the satellite a change in the particle flux is observed. As we will show, a similar process is responsible for the modulation of the observed particle flux in CTX. The particle motion in CTX is chaotic but the decorrelation time is long with respect to a drift period. Thus inhomogeneities in phase space propagate azimuthally and lead to temporal variations in the observed particle flux.

The Poincaré surface of section, while a useful way of obtaining topological information about phase space, does not provide sufficient information about the temporal evolution of particle trajectories to understand the modulation observed in the electron flux. The evolution of phase space, however, can be examined by computing the Hamiltonian phase space flow. For a $1\frac{1}{2}$ degree of freedom Hamiltonian, flows can be computed by dividing the (ψ,φ) plane into an $M\times M$ grid and evolving the equations of motion for each grid point. Since radial transport is of interest here, trajectories initially on the same radial flux surface are assigned the same color. In order to maintain a uniform distribution of grid points, the equations of motion are integrated backwards in time.

An example of a Hamiltonian phase space flow is shown in Fig. 8 where six "frames" from an animation of the phase space flow are displayed. The electrostatic potential used in the equations of motion is the same as was used to compute the Poincaré surface of section shown in Fig. 7. Clearly evident is the relatively slow, inhomogeneous mixing of phase space.

To understand how these pictures relate to the experimental observations, recall that the use of ECRH leads to a strong radial gradient in the hot electron density. The different colors used to identify a phase space element's initial radial position could also represent the initial particle density of the phase space element. Thus a consequence of the relatively slow mixing of phase space is that the density of particles reaching the detector has a very complicated time dependence. These simulations indicate that it is this process that is responsible for the modulation of the flux of electrons to the gridded particle detector observed in the experiment.

D. Hamiltonian transport simulation

In order to demonstrate the relationship between the chaotic mixing of phase space and the modulation of the energetic electron flux more directly, a numerical simulation which computes the time-evolution of the flux of energetic electrons to a small region of phase space has been performed. The objective of the simulation is to reproduce the gross frequency and amplitude of the modulation present in the observed energetic electron flux.

In the code, an ensemble of particle trajectories that are at the "gridded particle detector" at time $t=T$ are randomly selected. The values for ψ and φ are selected randomly from a region of phase space determined by the position and the area of the probe. As with the Poincaré surfaces of section and the Hamiltonian phase space flow, spectral information from the experiment is used to construct the electrostatic potential in the form of Eq. (4). However, in order to model the growth of the wave the spectrum is modified to include a time-dependent wave amplitude: $\Phi_0\rightarrow\Phi_0(t)=\Phi_0\tanh(\tau^2t^2)$, where τ is the growth time of the fluctuations. To increase computational efficiency, the equations of motion are integrated backwards in time from $t=T$ to $t=0$ rather than forward in time.

To model the finite extent of the experimental device, trajectories that move beyond the vacuum chamber wall, $R>R_w$, are considered lost. Particles may also be lost if they strike the metal column which extends from the wall to support the gridded particle detector or the detector itself. In the simulation, any trajectory which crosses the location of the probe, $R>R_d$ and $\varphi=\varphi_d$, has a finite probability of being removed.

At $t=0$ the probability that a trajectory came from a model hot electron distribution, $F_h(\mu,\psi)$, is computed. The hot electron distribution is assumed to be of the form:

$$F_h(\mu,\psi,t=0)=N_h\alpha(\psi,t=0)f(\mu), \tag{6}$$

where $N_h\equiv n_hV$ is the number of particles within a flux tube of volume V. The initial spatial dependence of the hot electron distribution is taken to be:

$$\alpha(\psi,t=0)=a(\psi/\psi_0-\psi_d/\psi_0)^b(2-\psi/\psi_0)^c, \tag{7}$$

where $\psi_0=B_0R_0^2$ and $\psi_d=B_0R_0^3/R_d$. Note that two of the three parameters a, b, and c are chosen so that $\alpha(\psi_0)=1$ and $\alpha'(\psi_0)=0$. The remaining parameter can be adjusted to change the steepness of the profile. The magnetic moment μ is chosen from an exponential distribution:

$$f(\mu)=\frac{1}{\mu_0}\exp(-\mu/\mu_0), \tag{8}$$

To model the energy resolution of the probe, only trajectories with $\mu B_0\in[E_{\min},E_{\max}]$ are integrated.

If a trajectory is determined to have come from the hot electron ring, it is used to construct the current:

$$I=en_hA\langle v\rangle=en_hA\frac{1}{\mathcal{N}}\sum_{i=1}^{\mathcal{N}}v_i, \tag{9}$$

where $\frac{1}{2}m_ev_i^2=\mu_iB(\psi_d)$ for the ith particle and A is the area of the probe. This process is repeated to construct $I(t)$.

The result of one such simulation is presented in Fig. 9 where it is compared with the corresponding energetic electron flux measurement from the experiment. The simulation was run with a randomly selected ensemble of 5000 particle trajectories. Also shown in Fig. 9 is the Fourier transform of the simulated and measured electrostatic fluctuations as well as the Fourier transform of the simulated and measured energetic electron flux. The simulation successfully reproduces the frequency of the observed modulations as well as the relative amplitude. Since the phases of the measured waves

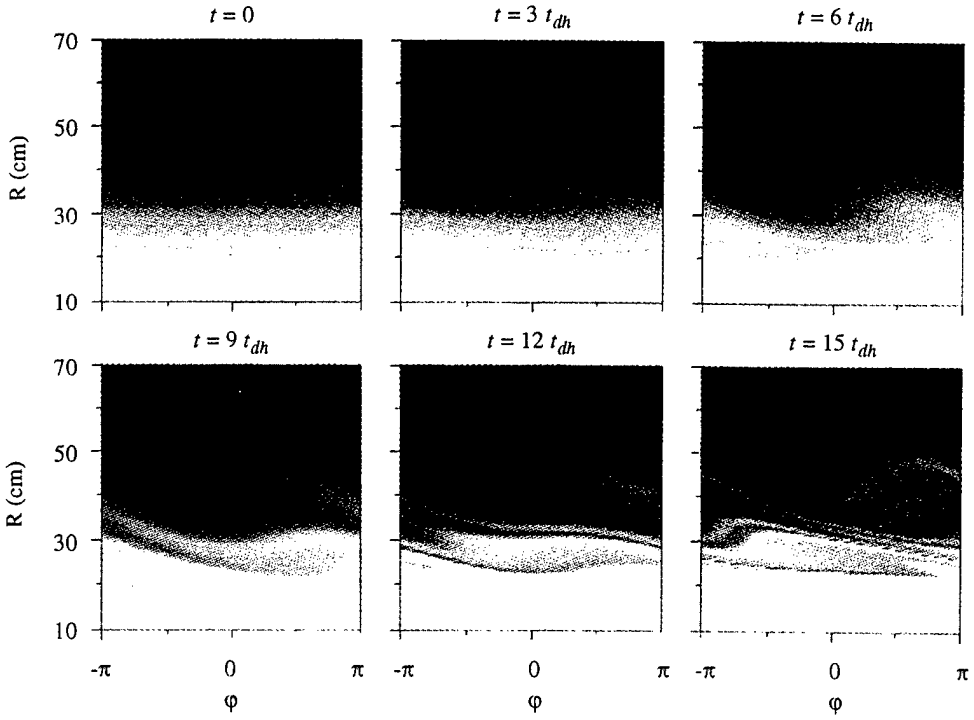

FIG. 8. Six "frames" from an animation of the Hamiltonian phase-space flow which show the slow, inhomogeneous mixing of phase space. The magnetic moment, μ, is chosen so that $\mu B_0 = 4$ keV ($t_{dh} = 3.34$ μs). $\tau = 10$ μs, $\Phi_0 = 150$ V and $M = 128$. Note that phase space is plotted in (R, φ) coordinates.

and other profile parameters are unknown, the simulation cannot be expected to reproduce the temporal evolution of the particle flux exactly.

E. Quasilinear transport simulation

In this section we explore the relationship between the Hamiltonian and quasilinear descriptions of chaotic particle transport. Quasilinear theory models transport as a random walk across flux surfaces with each step uncorrelated with the previous one. As is demonstrated by the phase space flows and the flux simulation, this does not describe the transport processes observed in the experiment. There are several reasons for this, the most important being that the measured fluctuation spectra have a finite bandwidth which allows particle trajectories to be correlated over several drift periods. Also, the fluctuation spectra may not lead to complete stochasticity. The presence of phase space structures, such as the islands shown in Fig. 7b, can increase the correlation time of particle trajectories.[11]

We have written a simple numerical simulation in order to compare the predictions of Hamiltonian and quasilinear descriptions of particle transport. In the code the simulated electron current is approximated by:

$$I(t) \approx eA \int d^3v v F_h(\mu, \psi_d, t) \sim eA n_0 \alpha(\psi_d, t) \langle v \rangle, \quad (10)$$

where $\langle v \rangle$ is the average velocity of particles reaching the detector. The function $\alpha(\psi, t)$ evolves according to a diffusion-like equation:

$$\frac{\partial \alpha}{\partial t} = \frac{\partial}{\partial \psi} \bar{D}_{\psi\psi} \frac{\partial \alpha}{\partial \psi} - \frac{\alpha}{t_L} \theta(\psi - \psi_d), \quad (11)$$

where $\bar{D}_{\psi\psi}$ is the diffusion coefficient averaged over resonant velocities. The second term in Eq. (11) models particle losses due to the probe and $\theta(\psi - \psi_d)$ is the unit step function. The boundary conditions are set so that the solution goes to zero at the wall and the magnet casing. The initial distribution for α is given by Eq. (7).

The quasilinear diffusion coefficient for the model fluctuation spectrum used in the Hamiltonian simulation [Eq. (4)] is given by:

$$D_{\psi\psi} = \frac{\pi m^2 c^2 \Phi_0^2}{2\mathscr{S}} \sum_{ml} |a_{ml}|^2 \delta(\omega_l - m\omega_{dh}(\mu, \psi)), \quad (12)$$

where $\delta(\omega_l - m\omega_{dh})$ is the Dirac delta function. The diffusion coefficient averaged over resonant velocities yields $\bar{D}_{\psi\psi}$. Alternatively, the diffusion coefficient can be determined from integrating the equations of motion as was done by Chen *et al.*[17] In this case the diffusion coefficient is computed using the formula:

$$\bar{D}_{\psi\psi} = \frac{\langle (\psi(t) - \psi(0))^2 \rangle}{2t} - \frac{\langle (\psi(t) - \psi(0)) \rangle^2}{2t}. \quad (13)$$

The agreement between these two approaches is generally good and the numerically determined value is used in the simulation.

The result of a single run representative of the quasilinear simulation is shown in Fig. 10 where it is compared with the result from the corresponding Hamiltonian simulation. Significantly, the predictions of the quasilinear model do not reproduce the experimentally observed modulations and underestimate the maximum particle fluxes measured in the experiment.

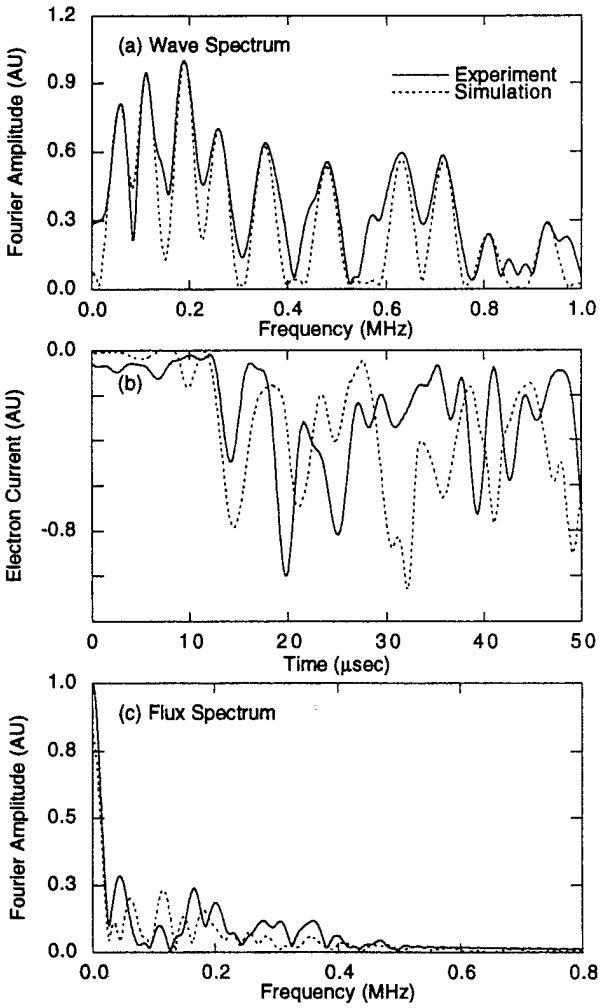

FIG. 9. A comparison of experiment and a simulation which reproduces the gross frequency and depth of the modulation. (a) Frequency spectrum of the electrostatic waves, (b) detected energetic electron flux, and (c) Fourier transform of the detected flux.

The failure of quasilinear theory to reproduce the salient features of the observed transport seems to be inconsistent with the strong chaos evident in the Poincaré surface of section shown in Fig. 7. We can reconcile this apparent discrepancy by considering the energetic electron flux predicted by the Hamiltonian simulation averaged over a magnetic flux surface. The result of such a calculation is presented in Fig. 10 where it is compared with the corresponding "point" result as well as the quasilinear result. The averaged energetic electron flux from the Hamiltonian simulation is very close to the quasilinear result indicating that the spatial resolution of the probe plays an important role in the deviations from quasilinear theory observed in the experiment.

V. SUMMARY

This paper reports the experimental observation of wave-induced chaotic radial transport of energetic electrons in a laboratory terrella. In the experiment ECRH is used to create a localized population of energetic, magnetically trapped electrons. The trapped electrons excite quasi-periodic, drift-resonant fluctuations which are identified as

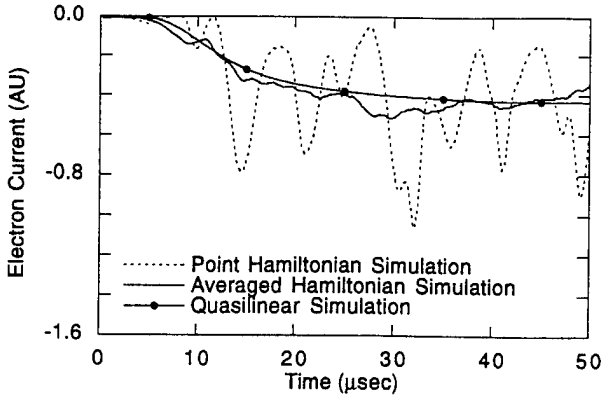

FIG. 10. A comparison of the energetic electron current computed from Hamiltonian and quasilinear models. The predictions of the quasilinear model lack the large-scale modulations seen in the experimental data and underestimate the maximum observed particle fluxes. There is, however, good agreement between the result of the Hamiltonian simulation averaged over a magnetic flux surface and the quasilinear simulation.

the hot electron interchange instability. Increases in the flux of energetic particles to a gridded particle detector are well correlated with the presence of the fluctuations. The measured particle flux is modulated at frequencies related to the precessional drift motion of the energetic electrons.

This paper has also presented the detailed results of numerical simulations based on the guiding center drift equations which model transport processes observed in the experiment. Most significantly, the simulations confirm that the observed transport is the result of nonlinear wave-particle interactions: transport is only observed when the wave-spectrum meets the conditions for global chaos; no transport is observed when the fluctuations lead to thin, radially localized bands of stochasticity.

The results of Hamiltonian simulations which explore the temporal details of the observed particle transport have also been presented. Animations of phase space flows have shown that the modulation of the flux of energetic electrons to the gridded particle detector can be understood in terms of the chaotic mixing of phase space. The relationship between chaotic mixing and the experimental measurements has also been demonstrated by a simulation which models the flux of energetic particles to a small region of phase space. The simulation reproduces the frequency and depth of the observed modulation. In contrast, quasilinear models of transport do not reproduce the modulation of the particle flux and underestimate the maximum particle fluxes observed in the experiment.

The results presented in this paper suggest that it may be worthwhile to model the temporal evolution of radiation belt particles in the magnetosphere using Hamiltonian methods. The results of such a study would better define the role that wave-induced chaotic radial transport plays in the dynamics of energetic particles in the magnetosphere. One candidate for driving drift-resonant chaotic radial transport similar to that observed in the experiment is low-m Pc 5 waves which have periods of 150–600 s. These ultra-low-frequency waves would drift-resonate with electrons with energies of 4–60

MeV at geosynchronous orbit and thus are strongly relativistic. Drift-resonant protons at geosynchronous orbit would have energies between 1–5 MeV but would be nonrelativistic.

Finally, we recognize that impulsive perturbations are of primary interest in magnetospheric transport of energetic particles but we stress that the modeling techniques used in this paper are clearly not restricted to wave-like perturbations. As mentioned in the introduction, Chen *et al.*[17] have modeled stormtime ring current proton enhancement due to impulsive variations in the Earth's convection electric field using the guiding center drift equations. They compared the dispersion averaged over a magnetic flux surface with the predictions of quasilinear theory and agreement between the two approaches was generally good. However, our results suggest that the averaged dispersion may not accurately represent the temporal evolution of energetic particle fluxes in localized regions of phase space.

ACKNOWLEDGMENTS

We would like to acknowledge the advice of Dr. Akira Hasegawa during the design of CTX and recognize the contributions of R. Singer and R. Maruyama to the implementation of the hard x-ray diagnostic. HPW would also like to thank R. Czerwinski for many helpful discussions.

This work has been supported by the Air Force Office of Scientific Research, Grant No. F4A-962093-10071, the National Aeronautics and Space Administration, Grant No. NAGW-3539, and the National Science Foundation, Grant No. ATM-91-11396.

[1]J. W. Dungey, Space Sci. Rev. **4**, 199 (1964).
[2]M. Schulz and L. J. Lanzerotti, *Particle Diffusion in the Radiation Belts, Physics and Chemistry in Space* (Springer-Verlag, New York, 1974), Vol. 7.
[3]A. L. Vampola, J. Electrostat. **20**, 3 (1987).
[4]D. N. Baker, J. B. Blake, L. B. Callis, J. R. Cummings, D. Hovestadt, S. Kanekal, B. Klecker, R. A. Mewaldt, and R. D. Zwickl, Geophys. Res. Lett. **21**, 409 (1994).
[5]R. M. Thorne, Pure Appl. Geophys. **118**, 128 (1980).
[6]M. P. Nakada and G. D. Mead, J. Geophys. Res. **70**, 4777 (1965).
[7]T. A. Farley, A. D. Tomassain, and M. Walt, Phys. Rev. Lett. **25**, 47 (1970).
[8]C. G. Fälthammer, J. Geophys. Res. **70**, 1487 (1965).
[9]W. H. Spjeldvik, J. Geophys. Res. **82**, 2801 (1977).
[10]R. B. Sheldon and D. C. Hamilton, J. Geophys. Res. **98**, 13491 (1993).
[11]G. M. Zaslavsky, D. Stevens, and H. Weitzner, Phys. Rev. E **48**, 1683 (1993).
[12]A. B. Rechester and R. B. White, Phys. Rev. Lett. **44**, 1586 (1980).
[13]R. G. Littlejohn, Phys. Fluids **20**, 2445 (1979).
[14]A. H. Boozer, Phys. Fluids **23**, 904 (1980).
[15]A. A. Chan, L. Chen, and R. B. White, Geophys. Res. Lett. **16**, 1133 (1989).
[16]M. W. Chen, M. Schulz, L. R. Lyons, and D. G. Gorney, Geophys. Res. Lett. **98**, 621 (1992).
[17]M. W. Chen, M. Schulz, L. R. Lyons, and D. G. Gorney, J. Geophys. Res. **98**, 3835 (1993).
[18]M. W. Chen, L. R. Lyons, and M. Schulz, J. Geophys. Res. **99**, 5745 (1994).
[19]H. P. Warren and M. E. Mauel, Phys. Rev. Lett. **74**, 1351 (1995).
[20]M. Mauel, H. P. Warren, and A. Hasegawa, IEEE Trans. Plasma Sci. **PS-20**, 626 (1992).
[21]N. A. Krall, Phys. Fluids **9**, 820 (1966).
[22]H. L. Berk, Phys. Fluids **19**, 1255 (1976).
[23]H. P. Warren, A. Bhattacharjee, and M. E. Mauel, Geophys. Res. Lett. **19**, 941 (1992).
[24]C. T. Hsu, C. Z. Cheng, P. Helander, D. J. Sigmar, and R. White, Phys. Rev. Lett. **72**, 2503 (1994).
[25]K. Schindler, Rev. Geophys. **7**, 51 (1969).
[26]V. D. Il'in and A. N. Il'ina, Sov. Phys. JETP **43**, 661 (1976).
[27]V. D. Il'in and A. N. Il'ina, Sov. Phys. JETP **45**, 514 (1977).
[28]V. Vasyliunas, J. Geophys. Res. **73**, 2839 (1968).
[29]R. C. Garner, M. E. Mauel, S. A. Hokin, R. S. Post, and D. L. Smatlak, Phys. Fluids B **2**, 242 (1990).
[30]S. H. Nawab and T. F. Quatieri, in *Advanced Topics in Signal Processing*, edited by J. S. Lim and A. V. Oppenheim (Prentice–Hall, Englewood Cliffs, NJ, 1988), pp. 289–337.
[31]S. Hiroe, J. B. Wilgin, F. W. Baity, L. A. Berry, R. J. Colchin, W. A. Davis, A. M. El Nadi, G. R. Haste, D. L. Hillis, D. A. Spong, and T. Uckan, Phys. Fluids **27**, 1019 (1984).
[32]M. J. Gerver and B. G. Lane, Phys. Fluids **29**, 2214 (1986).
[33]A. G. Kornienko, M. Y. Natenzon, R. Z. Sagdeev, and G. M. Zaslavsky, Phys. Lett. A **158**, 398 (1991).
[34]W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes: The Art of Scientific Computing* (Cambridge University Press, Cambridge, 1986).
[35]A. J. Lichtenberg and M. A. Lieberman, *Regular and Stochastic Motion*, Applied Mathematical Sciences (Springer-Verlag, New York, 1983).
[36]L. J. Lanzerotti, C. S. Roberts, and W. L. Brown, J. Geophys. Res. **72**, 5893 (1967).
[37]A. L. Vampola and A. Korth, Geophys. Res. Lett. **19**, 625 (1992).

Observation of wave-induced chaotic radial transport in a laboratory terrella experiment*

H. P. Warren,[†,a)] M. E. Mauel, D. Brennan, and S. Taromina
Department of Applied Physics, Columbia University, New York, New York 10027

(Received 8 November 1995; accepted 4 January 1996)

The wave-induced chaotic radial transport of energetic electrons has been observed in a laboratory terrella, the Collisionless Terrella Experiment (CTX) [H. P. Warren and M. E. Mauel, Phys. Plasmas **2**, 4185 (1995)]. In the experiment electron-cyclotron-resonance heating (ECRH) is used to create a localized population of trapped energetic electrons which excite drift-resonant electrostatic fluctuations. Measurements with multiple high-impedance floating potential probes are used to determine the amplitude, frequency, and mode structure of the observed fluctuations. Energetic electron transport is observed with gridded particle detectors. Poincaré surfaces of section indicate that increases in the flux of energetic electrons to the detectors occur only when fluctuations which meet the conditions for global chaos are present. Quasilinear transport simulations do not reproduce several important features of the experimental measurements. In contrast, Hamiltonian simulations reproduce many of the salient temporal characteristics of the experimental measurements and indicate that the persistence of phase-space correlations plays an important role in the energetic electron transport observed in the experiment. © *1996 American Institute of Physics.* [S1070-664X(96)91405-X]

I. INTRODUCTION

Understanding the transport of energetic particles is a fundamental problem of laboratory, astrophysical, and space plasmas. Of particular interest is transport in collisionless plasmas driven by non-linear wave-particle resonances. For example, chaotic radial transport which preserves the first two adiabatic invariants μ and J but breaks ψ is an essential mechanism for both the injection of particles into the Earth's magnetosphere and their acceleration.[1] This transport is driven by sudden, large-scale perturbations in the Earth's geomagnetic and convection electric fields caused by variations in the solar wind and interplanetary magnetic field.[2]

Previously we reported the first observation of wave-induced chaotic radial transport in a laboratory terrella experiment, the Collisionless Terrella Experiment (CTX).[3,4] One of the primary goals of CTX is to study the process of chaotic radial transport in dipole magnetic fields which preserves the first two adiabatic invariants μ and J.[5] In particular, these experiments investigate the relationship between fluctuation spectra and models of energetic particle transport and provide the first laboratory tests of Hamiltonian methods which can be used to simulate transport in collisionless plasmas.

In the experiment electron-cyclotron-resonance heating (ECRH) is used to create a highly localized population of magnetically trapped energetic electrons.[6] While the ECRH is on, the energetic electrons excite quasiperiodic bursts of electrostatic fluctuations which resonate with the precessional drift motion of the trapped particles. Drift-resonant electrostatic waves are also observed during the afterglow of the plasma discharge, when the ECRH has been turned off. During both times, the observed fluctuations have a complicated time-dependent frequency content.

By computing Poincaré surfaces of section we find that during the instability bursts the measured spectral characteristics of the drift-resonant fluctuations meet the conditions required for global chaotic particle transport. Coincident with the instability bursts, energetic electron transport is observed with gridded particle detectors located outside the hot electron ring. The observed transport is strongly modulated at frequencies related to the precessional drift-frequency of the energetic electrons. During the afterglow, Poincaré surfaces of section indicate that the instability wave spectrum does not satisfy the conditions for global chaos and no enhanced transport is observed experimentally.

The Poincaré surfaces of section which model particle motion during the instability bursts show that phase space is strongly chaotic and suggest that quasilinear theory should be applicable. However, in order to model the spatial and temporal evolution of particle fluxes observed in the experiment, simulations using the guiding center drift Hamiltonian are necessary. One such simulation involves computing the flux of equatorial particle trajectories to a small region of phase space which represents the particle detector. This simulation reproduces the modulation depth and frequency of the experimentally observed electron flux. When the results of the Hamiltonian transport simulation are compared with a simple quasilinear transport model we find that the predictions of the quasilinear model do not reproduce the experimentally observed modulations and underestimate the maximum particle fluxes measured in the experiment.

II. THE COLLISIONLESS TERRELLA EXPERIMENT

The CTX experimental device consists of a dipole electromagnet suspended mechanically in an axisymmetric aluminum vacuum vessel 1.4 m in diameter. The magnet is sup-

*Paper 2IA1, Bull. Am. Phys. Soc. **40**, 1666 (1995).
[†]Invited speaker.
[a)]Present address: Naval Research Laboratory, Washington, DC 20375.

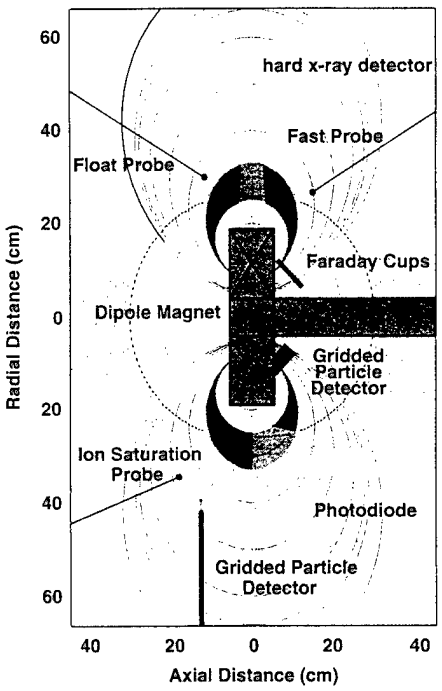

FIG. 1. The magnetic field geometry of CTX. The solid lines represent magnetic field lines and the dotted lines represent surfaces of constant magnetic field. The location of several plasma diagnostics and the approximate location of the hot electron ring are also shown.

ported by a stainless steel casing which houses the electrical and cooling leads. The strength of the dipole magnetic field is approximately 15 kG at the face of the magnet and falls off to less than 50 G at the wall. The magnetic field lines and mod-B surfaces of the dipole magnet used in CTX and are shown in Fig. 1.

A pulse of microwaves lasting approximately 0.5 s is used to break down and heat hydrogen gas which is puffed into the vacuum chamber both at the beginning and during a discharge. The microwave source is a continuous wave magnetron which has a peak output of 1.5 kW at a frequency of 2.45 GHz. The electron gyrofrequency equals the frequency of the applied microwaves along the surface defined by $B_0 \equiv B \simeq 875$ G. The equatorial field strength of the dipole magnet has this value at a radial distance of $R_0 \equiv R \simeq 27$ cm and this defines the center of the hot electron ring or artificial radiation belt.

Plasma diagnostics include a series of Langmuir and high impedance floating potential probes situated at five locations throughout the vacuum vessel. The probes can be repositioned radially to examine the density and potential fluctuations at different flux surfaces. Multiple probe measurements allow the direction of propagation, the azimuthal mode number and radial mode structure of the observed fluctuations to be determined.

A krypton proportional counter is positioned at the midplane of the magnetic field and is used to measure hard x-ray emission at energies between 1–60 keV. The counter is collimated to view the equatorial region of the dipole magnetic field. X-ray spectra are collected and stored at 50 ms intervals during the plasma discharge.

Gridded particle detectors are located inside the vacuum chamber both near the equatorial midplane and in the polar region. The equatorial particle detector is supported by a metal rod which extends from the vacuum chamber wall and can be repositioned to measure particle flux at different radial positions. There are five polar particle detectors which are situated so that the magnetic field at each detector is the same. The gridded particle detectors are biased to repel ions and electrons with energies less than 100 eV.

Five Faraday cups are also located in the polar region of the dipole magnetic field. The Faraday cups are situated to view the same field lines as the polar gridded particle detectors but rotated 90° in azimuthal angle. The Faraday cups are biased to repel ions.

Three photodiodes are positioned to measure radiation emitted from the plasma at wavelengths from 400–1100 nm. The photodiodes can be collimated to look at the inside of the hot electron ring ($R<27$ cm), the center of plasma (27 cm$<R<$54 cm), and the outside of the plasma ($R>54$ cm).

III. EXPERIMENTAL OBSERVATIONS

The intensity of the hot electron population is characterized by the hard x-ray emission produced by electron-ion and electron-neutral bremsstrahlung. A fraction of these x-rays are detected with the krypton proportional counter. The observed distributions are non-Maxwellian, characteristic of microwave-heated electrons.[7] The electrons with energies between 1–10 keV are referred to as the "warm" population, and electrons with energies above 10 keV are referred to as the "hot" population. When the microwave power is switched off, the "hot" population persists for 5–20 ms, defining the discharge "afterglow."

When an intense hot electron population is produced, drift-resonant fluctuations ($\omega \sim \omega_{dh}$) are observed both while the ECR heating is on and in the afterglow. During the heating, the fluctuations occur in quasiperiodic bursts lasting approximately 300–500 μs. During the afterglow, the drift-resonant oscillations persist for as long as several milliseconds. At both times, the observed instabilities propagate azimuthally in the direction of the electron ∇B drift, are flute-like with a constant phase along a field-line, and have a broad radial structure extending throughout the plasma. The saturated amplitudes of the floating potential oscillations present at both times are similar, typically 100–200 V.

Fourier analysis indicates that the quasiperiodic bursts consist of incoherent, broad-band fluctuations with frequencies typically below 2 MHz, although some particularly intense bursts have frequencies as high as 5 MHz. The instabilities observed during the afterglow are more coherent than those observed during the ECRH and have a higher range of frequencies, $f \sim$ 1–12 MHz.

The fluctuations observed during the ECRH also differ from those observed in the afterglow in terms of their azimuthal mode number. During the quasiperiodic instability bursts, the azimuthal mode number is usually limited to $m=1$, except in the most intense bursts when some high-frequency, $m=2$ modes are observed. In the afterglow,

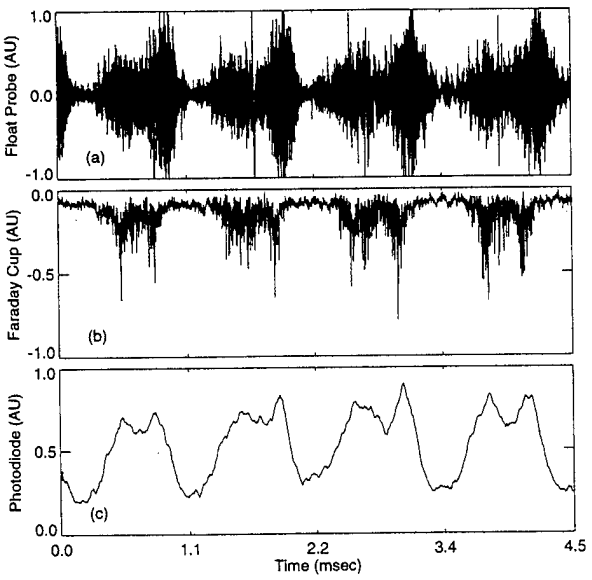

FIG. 2. Simultaneous measurements of drift-resonant fluctuations and energetic electron transport during the microwave heating. (a) Floating potential oscillations illustrate intense, quasiperiodic bursts correlated with (b) rapid increase of electron flux to the Faraday cup and (c) visible photon emission from the plasma measured by the center photodiode.

$m \leqslant 6$, and there are often multiple waves with the same azimuthal mode number present simultaneously.

The drift-resonant fluctuations observed in the experiment are related to the hot electron interchange instability that has been observed in other ECR heated magnetic mirror experiments.[8,9] The linear theory of the hot electron interchange, assuming slab geometry and a monoenergetic distribution of electrons, was developed by Krall[10] for $\omega \ll \omega_{ci}$ and by Berk[11] for $\omega \sim \omega_{ci}$ and the characteristics of the observed fluctuations are generally similar to those predicted by these simple theories.

Coincident with the bursts of wave activity observed during the ECRH are increases in the flux of energetic electrons to the equatorial gridded particle detector and the polar gridded particle detectors. As shown in Fig. 2, the radiation detected by the center photodiode also rises significantly during the instability bursts and is diminished between them. This last fact is particularly interesting. Recall that the use of microwave heating leads to a highly localized plasma. Beyond the hot electron ring the density of neutral hydrogen should increase rapidly. As the drift-resonant fluctuations drive energetic electrons out radially, they can ionize the surrounding neutral gas. Also recall that μ and J conserving radial transport leads to an adiabatic decompression with particle energy decreasing as $\sim 1/R^3$. This increases the Coulomb cross section and makes both electron-neutral and electron-ion collisions more frequent.

The characteristics of the transport measurements suggest a definite relationship between the spectral content of the waves and the observed energetic particle flux. The level of the electron flux is not simply related to the amplitude of the fluctuations but depends on the spectral content of the wave. Also note that the energetic electron flux observed both with the equatorial gridded particle detector and the Faraday cups is strongly modulated at frequencies related to the precessional drift-frequency electrons. The connection between the spectral content of the fluctuations and the electron flux will be examined in detail in the next section.

IV. WAVE-INDUCED PARTICLE TRANSPORT

The interaction of energetic electrons with the drift-resonant electrostatic waves observed in the experiment can be described by the guiding center drift Hamiltonian[12]:

$$\mathscr{H} = \frac{m_e c}{2e} \rho_{\|}^2 B^2 + \mu \frac{cB}{e} - c\Phi, \tag{1}$$

where m_e and e are the electron mass and charge, c is the speed of light, B is the magnitude of the dipole magnetic field, $\mu \equiv m v_{\perp}^2 / 2B$ is the magnetic moment, $\rho_{\|} \equiv v_{\|}/B$, and Φ is the electrostatic potential. For a curl-free magnetic field, the canonical coordinates of the guiding center drift Hamiltonian, $(\rho_{\|}, \chi)$, and (ψ, φ), are essentially the magnetic coordinates defined by: $\mathbf{B} = \nabla\psi \times \nabla\varphi = \nabla\chi$. The function $\psi = M \sin^2\theta/R$ is proportional to the magnetic flux bounded by a field line and the function $\chi = M \cos\theta/R^2$ is related to the distance along a field line. Note that $M \equiv B_0 R_0^3$ is the moment of the dipole magnet and (R, φ, θ) are spherical coordinates. Also note that the guiding center drift Hamiltonian is appropriate since energetic electrons produced in the experiment are non-relativistic.

When the particle motion is confined to the equatorial midplane of the magnetic field, $\rho_{\|} = \chi = 0$, $B = B(\psi)$, and the equations of motion are reduced to a particularly simple form. In general, particles will have a finite parallel velocity and will not be confined to the equatorial midplane. However, the precessional drift-frequency is only weakly dependent on pitch angle[1] and the results presented here would not be significantly altered by including parallel velocity.

The experimental observations described in the previous section indicate that the observed fluctuations can be modeled as a sum of traveling waves of the form:

$$\Phi(\varphi, t) = \frac{\Phi_0}{\sqrt{\mathscr{S}}} \sum_{m,l}^{N} a_{ml} \cos(m\varphi - \omega_l t + \varphi_l), \tag{2}$$

where $\mathscr{S} = \Sigma_{ml} |a_{ml}|^2$. The relative amplitudes, azimuthal mode numbers, frequencies, and phases are determined from the Fourier analysis of experimentally measured Langmuir probe signals. The waves observed in the experiment have a time-dependent frequency content, however, typically $t_{dh} d \log(a_{ml})/dt \ll 1$ and here we consider transport processes which occur on the time scale of several drift-periods.

For a single electrostatic wave the island half-width, $\Delta\psi$, at the resonant surface defined by $\omega_l - m\omega_{dh}(\mu, \psi_{lm}^r) = 0$, is given by: $\Delta\psi \simeq (2c\Phi_0 \psi_{lm}^r / \omega_{dh})^{1/2}$. When multiple drift-resonant waves with sufficiently large amplitude are present, islands will overlap and lead to chaotic transport. The topological properties of particle trajectories can be determined by constructing the Poincaré surface of section which is formed by plotting (ψ, φ) at multiples of the mapping time. For a spectrum of waves, the mapping time is the least common multiple of the periods of the wave motion: $T_M = \mathrm{LCM}(T_l = 2\pi/\omega_l)$.[13] In practice, the frequen-

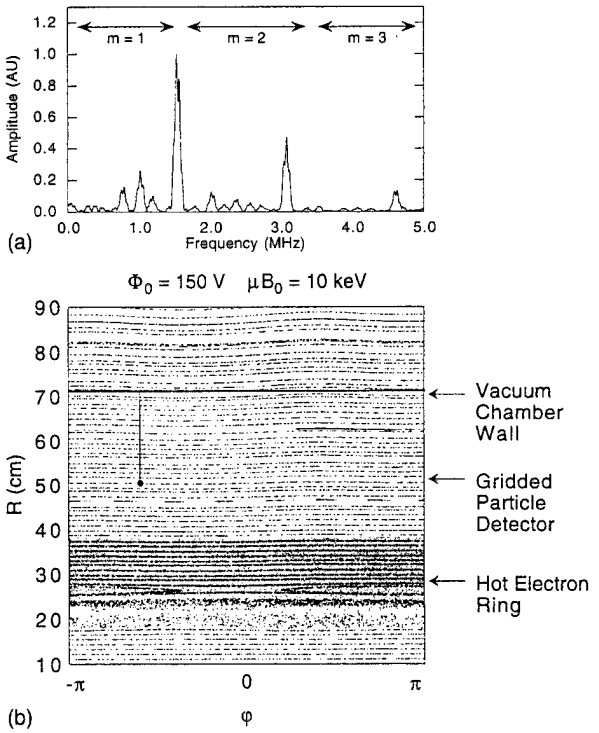

FIG. 3. A spectrum of waves taken from experimental measurements of fluctuations observed during the afterglow and a corresponding Poincaré surface of section. The chaos is localized to a thin, radially localized band near the hot electron ring and consistent with the absence of electron flux to the particle detector. The magnetic moment, μ, is chosen so that $\mu B_0 = 10$ keV; other parameters are $\Phi_0 = 150$ V, $N = 10$, and $T_M = 100$ μs. Note that the surface of section is plotted in (R, φ) coordinates.

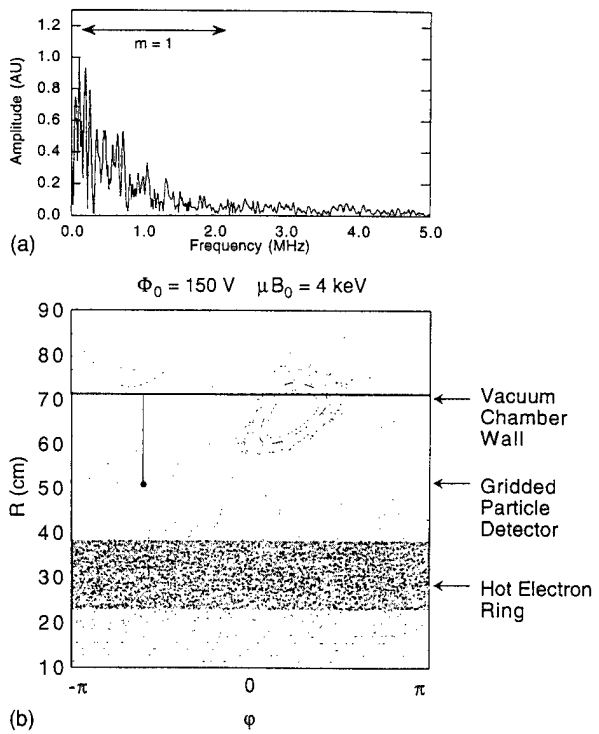

FIG. 4. A spectrum of waves taken from experimental measurements of fluctuations observed during the ECRH and a corresponding Poincaré surface of section. The fluctuations spectrum leads to chaos in ψ over the radial extent of the plasma. The magnetic moment, μ, is chosen so that $\mu B_0 = 4$ keV; other parameters are $\Phi_0 = 150$ V, $N = 10$, and $T_M = 100$ μs. Note that the surface of section is plotted in (R, φ) coordinates.

cies are rounded off to keep T_M from becoming too large. The equations of motion are integrated using the Gragg-Burlish-Stoer implementation of Richardson extrapolation.[14]

An example of a fluctuation spectrum used to model the electrostatic potential observed during the afterglow is shown in Fig. 3a. For electrons with energies above $\mu B_0 \approx 10$ keV this spectrum leads to resonant wave-particle interactions. However, as the Poincaré surface of section shown in Fig. 3b illustrates, the thin bands of chaos in ψ are limited to the proximity of the hot electron ring and do not extend out to the radial location of the gridded particle detector unless the amplitude of the potential fluctuations is very large.

A number of model fluctuation spectra have been constructed to model the potential fluctuations observed during the ECRH. For the wave spectrum shown in Fig. 4a, which is taken from the initial part of an instability burst, drift-resonances exist for electrons with energies between 1–10 keV (i.e., the "warm" electrons) from the center of the hot electron ring to the wall of the vacuum chamber. As shown in Fig. 4b, this spectrum of waves leads to chaos in ψ over the radial extent of the plasma. Examination of the phase space portraits indicates that when the wave amplitude is above $\Phi_0 \approx 75$ V, there are no encircling Kolmogorov-Arnol'd-Moser (KAM) surfaces[15] preventing global transport.

In order to demonstrate the relationship between the modulation of the current and the spectral content of the observed fluctuations, we have simulated the time-evolution of the flux of electrons to a small region of phase space which represents the particle detector. In the simulation, we randomly selected an ensemble of 5000 particle trajectories that are at the "gridded particle detector" at time $t = T$. Using spectral information from the experiment to construct the electrostatic potential in the form of Eq. (2), the equations of motion are integrated backwards in time from $t = T$ to $t = 0$ and the probability that the trajectory came from an assumed hot electron distribution, $F_h(\mu, \psi)$, is computed. This process is repeated in order to compute the current as a function of time.

The result of one such simulation is presented in Fig. 5 where it is compared with the corresponding energetic electron flux measurement from the experiment. Also shown in Fig. 5 is the Fourier transform of the simulated and measured electrostatic fluctuations as well as the Fourier transform of the simulated and measured energetic electron flux. The simulation successfully reproduces the frequency of the observed modulations as well as the relative amplitude. Since the phases of the measured waves and other profile parameters are unknown, the simulation cannot be expected to reproduce the temporal evolution of the particle flux exactly.

We have written a simple numerical simulation in order to compare the predictions of Hamiltonian and quasilinear descriptions of particle transport. In the code the simulated electron current is approximated by:

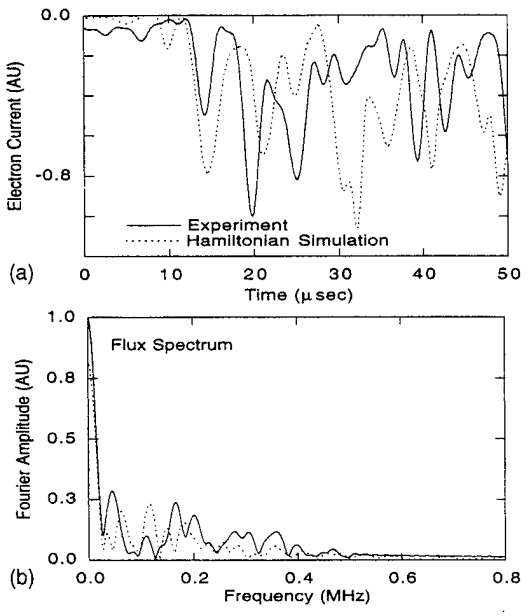

FIG. 5. A comparison of experiment and a simulation which reproduces the gross frequency and depth of the modulation. (a) Detected energetic electron flux, and (b) Fourier transform of the detected flux.

$$I(t)\approx eA\int d^3v\, v F_h(\mu,\psi_d,t)\sim eAn_0\alpha(\psi_d,t)\langle v\rangle, \qquad (3)$$

where $\langle v\rangle$ is the average velocity of particles reaching the detector. The function $\alpha(\psi,t)$ evolves according to a diffusion-like equation:

$$\frac{\partial\alpha}{\partial t}=\frac{\partial}{\partial\psi}\bar{D}_{\psi\psi}\frac{\partial\alpha}{\partial\psi}-\frac{\alpha}{t_L}\theta(\psi-\psi_d), \qquad (4)$$

where $\bar{D}_{\psi\psi}$ is the diffusion coefficient averaged over resonant velocities. The second term in Eq. (4) models particle losses due to the probe and $\theta(\psi-\psi_d)$ is the unit step function. The boundary conditions are set so that the solution goes to zero at the wall and the magnet casing.

The diffusion coefficient can be determined from integrating the equations of motion. In this case the diffusion coefficient is computed using the formula:

$$\bar{D}_{\psi\psi}=\frac{\langle(\psi(t)-\psi(0))^2\rangle}{2t} \qquad (5)$$

and the numerically determined value is used in the quasilinear simulation.

The result of a single run representative of the quasilinear simulation is shown in Fig. 6 where it is compared with the result from the corresponding Hamiltonian simulation. Significantly, the predictions of the quasilinear model do not reproduce the experimentally observed modulations and underestimate the maximum particle fluxes measured in the experiment.

V. SUMMARY AND DISCUSSION

This paper describes the experimental observation of wave-induced chaotic radial transport of energetic electrons in a laboratory terrella. In the experiment ECRH is used to

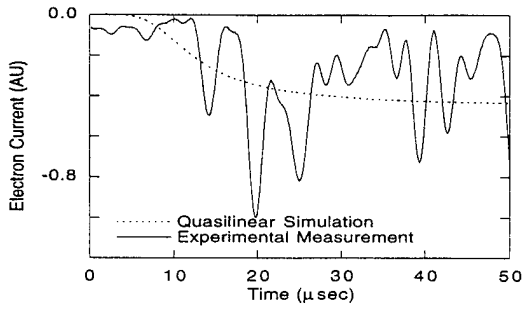

FIG. 6. A comparison of the experimental observation with the predictions of a quasilinear transport simulation. The predictions of the quasilinear model lack the large-scale modulations seen in the experimental data and underestimate that maximum observed particle fluxes.

create a localized population of energetic, magnetically trapped electrons. The trapped electrons excite quasiperiodic, drift-resonant fluctuations which are identified as the hot electron interchange instability. Increases in the flux of energetic particles to both the gridded particle detector situated near the equatorial midplane and the array of particle detectors located in the polar region are well correlated with the presence of the fluctuations. The radiation detected by the center photodiode also rises significantly during the instability bursts and is diminished between them. The measured particle fluxes are modulated at frequencies related to the precessional drift motion of the energetic electrons.

This paper has also presented the detailed results of numerical simulations based on the guiding center drift equations which model transport processes observed in the experiment. Most significantly, the simulations confirm that the observed transport is the result of non-linear wave-particle interactions: transport is only observed when the wave-spectrum meets the conditions for global chaos; no transport is observed when the fluctuations lead to thin, radially localized bands of stochasticity.

The results of Hamiltonian simulations which explore the temporal details of the observed particle transport have also been presented. The relationship between chaotic mixing and the experimental measurements has been demonstrated by a simulation which models the flux of energetic particles to a small region of phase space. The simulation reproduces the frequency and depth of the observed modulation. In contrast, quasilinear models of transport do not reproduce the modulation of the particle flux and underestimate the maximum particle fluxes observed in the experiment.

The simulations described here model transport processes which occur on the time scale of several drift-periods. This time scale is short when compared with the evolution of an entire instability burst. Future research will focus on coupling transport simulations to a dispersion relation for the instability in order to investigate the self-consistent, nonlinear evolution of the plasma. This will allow us to address some of the unresolved questions related to the saturation of the instability and the chirping of the wave frequeneies.

ACKNOWLEDGMENTS

This work has been supported by Air Force Office of Scientific Research Grant No. F4A-962093-10071, National Aeronautics and Space Administration Grant No. NAGW-3539, and National Science Foundation Grant No. ATM-91-11396.

[1]M. Schulz and L. J. Lanzerotti, *Particle Diffusion in the Radiation Belts*, No. 7 in *Physics and Chemistry in Space* (Springer-Verlag, New York, 1974).
[2]J. W. Dungey, Space Sci. Rev. **4**, 199 (1964).
[3]H. P. Warren and M. E. Mauel, Phys. Rev. Lett. **74**, 1351 (1995).
[4]H. P. Warren and M. E. Mauel, Phys. Plasmas **2**, 4185 (1995).
[5]H. P. Warren, A. Bhattacharjee, and M. E. Mauel, Geophys. Res. Lett. **19**, 941 (1992).
[6]M. Mauel, H. P. Warren, and A. Hasegawa, IEEE Trans. Plasma Sci. **PS-20**, 626 (1992).
[7]For example, R. A. Dandl, A. C. England, W. B. Ard, H. O. Eason, M. C. Becker, and G. M. Haas, Nuc. Fusion **4**, 344 (1964) or H. Ikegami, H. Ikezi, M. Hosokawa, S. Tanaka, and K. Takayama, Phys. Rev. Lett. **19**, 778 (1967).
[8]S. Hiroe, J. B. Wilgin, F. W. Baity, L. A. Berry, R. J. Colchin, W. A. Davis, A. M. El Nadi, G. R. Haste, D. L. Hillis, D. A. Spong, and T. Uckan, Phys. Fluids **27**, 1019 (1984).
[9]M. J. Gerver and B. G. Lane, Phys. Fluids **29**, 2214 (1986).
[10]N. A. Krall, Phys. Fluids **9**, 820 (1966).
[11]H. L. Berk, Phys. Fluids **19**, 1255 (1976).
[12]A. H. Boozer, Phys. Fluids **23**, 904 (1980).
[13]A. G. Kornienko, M. Y. Natenzon, R. Z. Sagdeev, and G. M. Zaslavsky, Phys. Lett. A **158**, 398 (1991).
[14]W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes: The Art of Scientific Computing* (Cambridge University Press, Cambridge, 1986).
[15]A. J. Lichtenberg and M. A. Lieberman, *Regular and Stochastic Motion*, No. 38 in *Applied Mathematical Sciences* (Springer-Verlag, New York, 1983).

Final Report on Subcontract with Columbia University under Grant No. F49620-93-10071 (1993-97) with the Air Force Office of Scientific Research

TITLE OF RESEARCH PROGRAM: Collisionless Dynamics of the Magnetosphere

SUBMITTED BY: Amitava Bhattacharjee, Professor
Department of Physics and Astronomy
University of Iowa, Iowa City, IA 52242

In our original proposal, we had proposed an extensive theoretical research program on substorms in the Earth's magnetosphere. As discussed below in our synopsis, and elaborated in the attached publications, we have made significant progress in developing a predictive dynamical theory of substorms. This theory has been obtained by a combination of analytical techniques and high-resolution computer simulation. Apart from the work on magnetospheric substorms, we have also carried out, supported by this grant (as well as the National Science Foundation), some interesting applications of our basic forced reconnection theory to laboratory plasmas. We have also completed a substantial body of work on current sheets and MHD turbulence in the solar corona. We were led into this research on the solar corona by the realization that there is much in common between thin current sheets in the magnetosphere and the solar corona, although the plasma parameters are quite different. Since current sheets mediate the transition to turbulence, it was natural then to investigate the spectral characteristics of the turbulence. In turn, these investigations led us to the investigation of vortex singularities and spectra in hydrodynamic turbulence. Although such diverse lines of investigation were neither envisioned nor proposed in our original proposal, they were an interesting and natural outgrowth of our effort on magnetospheric physics.

In what follows, we give a synopsis of our principal findings in magnetospheric (Section I) and solar physics (Section II). Details are described in the attached publications. This research supported by the AFOSR has led to 17 archival publications and 12 invited/selected talks in national/international conferences/workshops.

I. GEOMETRY AND DYNAMICS OF MAGNETIC RECONNECTION: SUBSTORMS IN THE MAGNETOSPHERE AND SOME LABORATORY APPLICATIONS

Magnetic reconnection can be free or forced. Free reconnection is caused by a spontaneous instability, whereas forced reconnection can occur in a stable plasma due to perturbations imposed on the boundary. We have investigated the relevance of both types of reconnection to the Earth's magnetotail.

Reconnection, if it is to be distinguished from diffusion, occurs in the vicinity of separatrices (such as neutral lines or X-points). Whereas separatrices are determined by the geometry of the magnetic field, the dynamics of reconnection is controlled by non-ideal mechanisms in a generalized Ohm's law such as resistivity, finite particle inertia or electron pressure tensor. If the geometry is unfavorable to topological change, it is not possible to trigger tearing instabilities by mere inclusion of non-ideal mechanisms in Ohm's law. This simple point is often misunderstood and has been a source of considerable controversy in the literature. The importance of geometry is illustrated by demonstrating the linear stability of the magnetotail in the presence of a significant B_z-field to all forms of tearing, independent of the mechanism that breaks field-lines/1,2/.

In the absence of free reconnection in the ambient magnetotail, attention is then

focused on the role of forced reconnection, driven by the solar wind. Multi-satellite observations indicate the development of thin current and a rapid intensification of the cross-tail current density at near-Earth distances during a short interval (<1 min) just before onset, after a period of sluggish growth (~0.5-1.5 hr). These observational features have been reproduced recently by analysis and MHD simulation, especially in the high-Lundquist-number regime/3,4,5/. In the slow growth and impulsive pre-onset phase, a thin current sheet develops spanning Y-points that stretch from the mid-tail region (~30 R_E) to the near-Earth region (~10 R_E). It is shown that a thin current sheet geometry involving Y-points and the associated convection pattern allows a natural resolution of some ongoing controversies in substorm research. When the current sheet becomes sufficiently thin, it is no longer justifiable to rely entirely on resistive MHD, and finite ion-Larmor-radius terms such as electron pressure gradients and Hall currents must be included in the theory. These collisionless effects (incorporated via the generalized Ohm's law) are shown to have a striking effect on thin current sheet and reconnection dynamics in the impulsive pre-onset phase/5/.

One of the important ground-based observational features of substorm dynamics is the evolution of the ionospheric potential in the auroral oval. We have proposed a kinematic, 3D reconnection model that integrates the geometry of the dayside and the nightside magnetosphere, and compares well observed ionospheric potential patterns/6/.

The basic principles of our work on nonlinear reconnection and the persistence of current sheets are also applicable to the important problems of sawtooth oscillations and mode-locking in laboratory fusion plasmas. These theoretical models are developed in refs. 7, 8 and 9. The boundary conditions and plasma regimes for laboratory applications are quite different than they are in the magnetosphere, and yet it is interesting to see that the same basic principles of free and forced reconnection are applicable to both types of plasmas.

II. SOLAR CORONAL DYNAMICS

One of the primary objectives of this research is the investigation of the effect of a small but finite resistivity on the dynamics of current sheet formation in the solar corona. It is demonstrated that helicity-conserving (free as well as forced) reconnection processes lead to the formation of current sheets near separatrices of the magnetic field. The rate of growth of the current sheet amplitude in the linear as well as the nonlinear regime of free and forced reconnection processes is calculated explicitly for some analytically tractable and realistic physical models/10,11/. These analytical results are compared with numerical simulations as well as observations. It is found that the heating produced by these current sheets is enough to account for the power balance in all quiet and some (but not all) active coronal loops.

We have continued the investigation of the time-history of current sheet formation and turbulence in 2D coronal equilibria. In the process, we have found some new and interesting results along some unexpected avenues of research. These include (i) the subtle role of the spatial extent of photospheric footpoint shear in controlling the impulsive growth of current sheets and reconnection /12/ (ii) the role of nonlinear shear-Alfven dynamics in determining the spectral character and anisotropy of MHD turbulence in the solar wind and the interstellar medium /13,14/ as well as the dynamo effect/15/, and (iii) new and rigorous results on the formation of finite-time vortex singularities in hydrodynamic flows (which are analogous to current singularities in MHD)/16,17/.

ARCHIVAL PUBLICATIONS SUPPORTED BY THE GRANT

1. Global asymptotic equilibria and collisionless tearing stability of magnetotail plasmas, X. Wang and A. Bhattacharjee, Journal of Geophysical Research **98**, 19419 (1993).

2. Tearing stability of the two-dimensional magnetotail, B. G. Harrold, A. Bhattacharjee, and X. Wang, Physics of Plasmas **2**, 3857 (1995).

3. Growth, sudden enhancement, and relaxation of current sheets in the magnetotail : two-dimensional substorm dynamics, Z. W. Ma, X. Wang and A. Bhattacharjee, Geophysical Research Letters **22**, 2985 (1995).

4. Fast reconnection and sudden enhancement of current sheets due to inward boundary flows, X. Wang, Z. W. Ma, and A. Bhattacharjee , Physics of Plasmas **3**, 2129 (1996).

5. Fast impulsive reconnection and current sheet intensification due to electron pressure gradients in semi-collisional plasmas, Z. W. Ma and A. Bhattacharjee, Geophysical Research Letters **23**, 1673 (1996).

6. A three-dimensional kinematic reconnection model of the magnetosphere: geometry and kinematics, Xiaogang Wang and A. Bhattacharjee, Journal of Geophysical Research **101**, 2641 (1996).

7. Nonlinear dynamics of m=1 kink-tearing modes in a modified magnetohydrodynamic model, X. Wang and A. Bhattacharjee, Physics of Plasmas **2**, 171 (1995).

8. Forced magnetic reconnection and the persistence of current sheets in static and rotating plasmas due to a sinusoidal boundary perturbation, Z. W. Ma, X. Wang, and A. Bhattacharjee, Physics of Plasmas **3**, 2427 (1996).

9. Forced reconnection and mode-locking in cylindrical plasmas, X. Wang and A. Bhattacharjee, Physics of Plasmas **4**, 748 (1997).

10. Current sheets and reconnection driven by footpoint motion in two-dimensional coronal loops with *X*-type neutral lines, X. Wang and A. Bhattacharjee, The Astrophysical Journal **420**, 415 (1994).

11. Dynamics of current sheet formation and reconnection in two-dimensional coronal loops, Z. W. Ma, C. S. Ng, X. Wang, and A. Bhattacharjee, Physics of Plasmas **2**, 3184 (1995).

12. Role of photospheric footpoint shear in the impulsive dynamics of the solar corona, Z. W. Ma and A. Bhattacharjee, Geophysical Research Letters **23**, 2955 (1996).

13. Self-consistency constraints on the dynamo mechanism, A. Bhattacharjee and Y. Yuan, The Astrophysical Journal **449**, 739 (1995).

14. Interaction of shear-Alfvén wave packets: implications for weak MHD turbulence in astrophysical plasmas, C. S. Ng and A. Bhattacharjee, The Astrophysical Journal **465**, 845 (1996).

15. Scaling of anisotropic energy spectra due to weak interaction of shear-Alfvén wave packets, C. S. Ng and A. Bhattacharjee, Physics of Plasmas **4**, 605 (1997).

16. Finite-time vortex singularity and Kolmogorov spectrum in a symmetric three-dimensional spiral model, A. Bhattacharjee, C. S.. Ng and X. Wang, Physical Review E **52**, 5110 (1995).

17. Sufficient condition for a finite-time singularity in a high-symmetry Euler flow: Analysis and statistics, C. S. Ng and A. Bhattacharjee, Physical Review E **54**, 1530 (1996).

INVITED/SELECTED TALKS IN CONFERENCE/WORKSHOPS

1. Global asymptotic equilibria and collisionless tearing stability of magnetotail plasmas, Mini-Workshop on Substorms, Johns Hopkins University Applied Physics Laboratory, March 22, 1993. (Collaborator : X. Wang)

2. Current and vortex singularities, Elizabeth and Frederick White Workshop on Fundamental Problems in the Physics of Magnetically Confined Plasmas, Australian Academy of Sciences, Canberra, 12-15 July, 1993. (Collaborator : X. Wang)

3. Fast sawtooth collapse in tokamaks, Elizabeth and Frederick White Workshop on Fundamental Problems in the Physics of Magnetically Confined Plasmas, Australian Academy of Sciences, Canberra, 12-15 July, 1993. (Collaborator : X. Wang)

4. Separatrices, singularities and turbulence, 1993 Cambridge Workshop on the Physics of Space Plasmas, Massachusetts Institute of Technology, July 19-23, 1993.

5. Collisionless tearing instabilities and their possible relevance for substorms, The International Conference on Substorms ICS-2, University of Alaska, Fairbanks, March 7-11, 1994. (Collaborator : X. Wang)

6. Role of current sheets in nonlinear magnetic island dynamics, *Theory of Fusion plasmas*, Joint Varenna-Lausanne International Workshop, International School of Plasma Physics <<Piero Caldirola>>, Varenna, August 22-26, 1994.

7. Geometry and dynamics of free and forced magnetic reconnection in the magnetosphere, 1994 Cambridge Workshop on the Physics of Space Plasmas, Bermuda, February 20-25, 1995. (Collaborators : Z.-W. Ma and X. Wang)

8. Separatrices, singularities and the Kolmogorov spectrum, March Meeting of the American Physical Society, San Jose, California, 20-24 March, 1995. (Collaborators: C. S. Ng and X. Wang)

9. Geometry and dynamics of reconnection: implications for substorms, Tutorial Lecture, Plenary Session of the GEM Workshop, Snowmass, Colorado, June 26-30, 1995.

10. Lectures on magnetic reconnection, SERC School on Magnetic Reconnection, Institute for Plasma Research, Ahmedabad, India, Decemeber 10-30, 1995.

11. Lectures on current sheets and vortex singularities, Nice School on "Vortex Tubes: Observations, Stability, Topology," Observatoire de la Côte d'Azur, Nice, France, May 19-25, 1996.

12. Geometry and dynamics of magnetic reconnection: implications for magnetotail plasmas, The Fifth International School/Symposium for Space Simulations, Radio Atmospheric Science Center, Kyoto University, Kyoto, Japan, March 13-19, 1997.

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 98, NO. A11, PAGES 19,419–19,434, NOVEMBER 1, 1993

Global Asymptotic Equilibria and Collisionless Tearing Stability of Magnetotail Plasmas

XIAOGANG WANG AND A. BHATTACHARJEE

Department of Applied Physics, Columbia University, New York

Asymptotic tail equilibria which are slowly varying in the Earth-Sun direction are constructed, including all three components of the magnetic field. These equilibria allow for spatial dependencies in B_n and B_y. Some of these equilibria are analyzed for stability with respect to collisionless electron tearing modes using a fluid model which predicts, to within a numerical factor of $\pi^{1/2}$, the growth rates derived from kinetic theory. No ion tearing instability is found. The B_y field is demonstrated to have a destabilizing effect on electron tearing modes. In the asymptotic equilibria considered here, electron tearing modes can grow in the presence of B_y in those regions where the stabilizing effect of electron bounce is small. Implications for numerical simulations and observations are discussed.

1. Introduction

Ever since Ness [1965] reported observational evidence for a neutral sheet in the Earth's magnetotail, the collisionless tearing instability has claimed much attention as a possible mechanism for magnetic reconnection in the tail. *Coppi et al.* [1966] considered a simple neutral sheet in which oppositely directed magnetic fields $B_x = B_x(z)$ face each other across the $z = 0$ line (in the $x - z$ plane) and demonstrated that such a sheet is unstable to the collisionless tearing instability [*Furth*, 1962; *Laval et al.*, 1966]. (We use here the standard solar magnetospheric coordinates (x,y,z), with the x axis in the Earth-Sun direction, the z axis in the south-north direction, and the y axis , which defines an ignorable direction, is chosen to make the coordinate system right-handed.) In the simple magnetic geometry considered by Coppi et al., the neutral line $z = 0$ is the source of the separatrix. Far away from the separatrix, the plasma obeys the ideal magnetohydrodynamic (MHD) equations. The departure from ideal MHD behavior occurs in a narrow region near the separatrix. By considering the energetics of the instability, Coppi et al. demonstrated that the dominant contribution to the inverse Landau damping effect comes from electrons, not ions. Thus this mode came to be known as the "electron tearing" instability.

We consider now the effect of a large, constant B_y field superimposed on the model of Coppi et al. The presence of B_y introduces magnetic shear in the model. The separatrix in the $x - z$ plane grows out of the $z = 0$ line. *Drake and Lee* [1977] showed that collisionless tearing modes are unstable in this geometry. (Strictly speaking, the results of *Drake and Lee* [1977] hold for a low-beta plasma without temperature gradients, as shown by *Cowley et al.* [1986]). Since the electrons carry the perturbed parallel current near the separatrix and provide the mechanism for reconnection through their small but finite inertia, this instability too can be classified as electron tearing.

Neither of the two cases discussed above are representative of the Earth's magnetotail. Much emphasis has been placed in the literature on the two-dimensional model

$$\mathbf{B} = B_0 \tanh\frac{z}{\lambda}\hat{x} + B_n \hat{z}, \qquad (1)$$

with constant B_0, λ, and B_n [*Schindler*, 1974; *Galeev and Zelenyi*, 1976; *Lembege and Pellat*, 1982; *Büchner and Zelenyi*, 1987; *Büchner et al.*, 1991; *Pellat et al.*, 1991; *Kuznetsova and Zelenyi*, 1991]. For nonzero values of B_n , this configuration has no magnetic separatrix. (For useful discussions of the role of separatrices in magnetic reconnection, we refer the reader to *Greene* [1988] and *Lau and Finn* [1990].) The absence of a separatrix implies that reconnection or tearing (in the sense of affecting a topological change) cannot really happen for significant values of B_n. It is widely believed that the "ion tearing" instability can occur in these circumstances, but the subject remains a matter of lively debate [*Schindler*, 1974; *Galeev and Zelenyi*, 1976; *Coroniti*, 1980; *Lembege and Pellat*, 1982; *Büchner and Zelenyi*, 1987; *Pellat et al.*, 1991; *Kuznetsova and Zelenyi*, 1991]. The analysis given in this paper turns out to support the point of view recently expressed by *Pellat et al.* [1991], who have questioned the existence of the ion tearing mode. This point of view has significant implications for electromagnetic particle simulations of collisionless tearing [*Terasawa*, 1981; *Hamilton and Eastwood*, 1982; *Swift*, 1983; *Ambrosiano et al.*, 1986; *Swift and Allen*, 1987; *Pritchett et al.*, 1989; *Zwingmann et al.*, 1990; *Pritchett et al.*, 1991] which can shed valuable light on this controversial issue.

Recently, we showed that the inclusion of a constant B_y field in the model (1) can qualitatively change its stability properties [*Wang et al.*, 1990; hereafter WBL]. Our model, referred to here as the three-component model, breaks with the tradition of using two-component models in theoretical analyses of collisionless instabilities in the magnetotail. However, just as in the two-component model, the formation of a magnetic separatrix is inhibited by the presence of a significant B_n field. It is therefore not surprising that we found that the electron tearing instability has a stabler parameter space and is much harder to excite in the three-component magnetotail than in a configuration with $B_n = 0$. Furthermore, the growth rate of the mode is slow, consistent with the growth phase, but not the expansion phase of a substorm.

We now develop an equilibrium model which is more realistic than considered heretofore in analytical studies of collisionless tearing. As in WBL, we include the B_y field. Fairfield and others have noted that the B_y field is a persistent feature of the magnetotail [*Fairfield*, 1979; *Cattell and Mozer*, 1982; *Lui*, 1983; *McComas et al.*, 1986; *Tsurutani et al.*, 1984; *Sibeck et al.*, 1985]. Voigt and coworkers [*Voigt and Hilmer*, 1987; *Hau and Voigt*, 1992] have shown, based on some analytical and numerical examples, that the requirements of global magnetostatic

Copyright 1993 by the American Geophysical Union.

Paper number 93JA01153.
0148-0227/93/93JA-01153$05.00

equilibrium of the magnetotail should include B_y. The possible role of the B_y field in observations of substorm dynamics was pointed out as early as 1978 by Akasofu and coworkers [*Akasofu et al.*, 1978]. From a detailed examination of IMP data, Akasofu et al. stressed the need for a three-component analysis of magnetic fields in the magnetotail during substorms. Though there are a number of other events reported in the literature in which an enhancement in B_y and a reduction in B_n is observed prior to the onset of a substorm [*Nishida et al.*, 1983; *Bieber et al.*, 1984; *Lepping*, 1987; *Takahashi et al.*, 1987; *Lui et al.*, 1988; *Lopez et al.*, 1989], no systematic studies of substorm events with correlated variations in B_y and B_n are available yet.

One of the main improvements of the present paper over WBL is the development of asymptotic equilibria with spatial dependencies in B_n and B_y, that is, $B_n = B_n(x,z)$ and $B_y = B_y(x,z)$. These equilibria, and the single-particle motions in them, are described in section 2. We show that these equilibria change qualitatively our current understanding of collisionless instabilities in the magnetotail by introducing new global features not captured adequately by the model,

$$\mathbf{B} = B_0 \tanh\frac{z}{\lambda}\hat{x} + B_y\hat{y} + B_n\hat{z}, \qquad (2)$$

with B_y and B_n constant. In this geometry, the global bounce period τ_{be} for electrons is much shorter than the growth time of the instability when $B_y \sim B_z$. WBL has been criticized for neglecting the stabilizing effect of this bounce [*Büchner et al.*, 1991; *Pritchett et al.*, 1991]. This criticism would be justified except for the fact that the model (2) is itself globally rather crude and underestimates severely the bounce period in a stretched magnetotail. If we must include the effect of electron bounce, then it is preferable to do so in an equilibrium model which captures the global features of the magnetotail with greater realism than equation (2). And that is precisely what we achieve by allowing for spatially varying B_n in the new equilibrium model. We then show that there are regions where $\gamma\tau_{be} \gtrsim 1$, and the electron tearing mode grows, with a growth rate γ, as predicted by WBL. Furthermore, when we consider the special case $B_n = \text{const}$ in which $\gamma\tau_{be} << 1$, we find no ion tearing, contrary to the findings of *Büchner et al.* [1991]. The persistence of slow electron tearing, and the absence of ion tearing, are recurrent themes that are explored in detail for both two-component ($B_y = 0$) and three-component ($B_y \neq 0$) equilibria in sections 4 and 5.

Though the electron tearing instability can account for the growth phase of a substorm, it is not sufficiently rapid to account for the current disruption and diversion that occurs at the onset of the expansion phase [*Takahashi et al.*, 1987; *Lui et al.*, 1988]. Elsewhere [*Wang et al.*, 1991], we have discussed that nonlinear mode-coupling effects may lead to a significant enhancement of the linear growth rate. It was implicitly assumed in that discussion that a linearly unstable mode will grow to sufficiently large amplitude that it can couple to other unstable modes. However, that possibility was explored in the context of the equilibrium represented by (2) which, as discussed here, has certain limitations. We suggest here that the circumstances in which collisionless tearing can grow as a robust instability, not only in the linear regime but also in the nonlinear regime, must involve reduction of B_n to extremely small values. We therefore point to the possibilities inherent in the three-component asymptotic equilibria calculated in this paper. An interesting property of these equilibria is that B_n may be reduced to zero at near-Earth distances when B_y is space-dependent. This possibility was first noted by *Hau and Voigt* [1992] whose profiles for B_y were different from ours, and who concluded that for their class of profiles, the B_y value required to reduce B_n to zero is much larger than the average value of B_y in the plasma sheet. For our class of profiles, we find that the reduction of B_n to zero occurs for values of B_y more in accord with observed values. When B_n vanishes, a separatrix can be formed, and tearing instabilities can grow. Whether the nonlinear evolution of these instabilities can actually account for the rapid current disruption and diversion observed in near-Earth regions remains an open question.

The layout of this paper is as follows. In section 2 we obtain some asymptotic tail equilibria, both with and without B_y. In section 3 we develop a fluid model, including a generalized Ohm's law, which allows the treatment of collisionless tearing modes, and benchmark the predictions of this model with known results from kinetic theory. In section 4 we use the energy integral derived from our fluid model to analyze the stability of the two-dimensional magnetotail without B_y and constant B_n. We find that the ion tearing mode does not occur, and the only possible instability, under certain conditions, is an electron tearing mode. In section 5 it is shown that the inclusion of B_y and a spatially varying B_n can further destabilize the electron tearing instability. We conclude in section 6 with a summary of our results and their implications for observations.

2. Some Asymptotic Equilibria

2.1. *Two-Component Equilibria*

We begin by considering some asymptotic two-component equilibria, i.e., equilibria for which $B_y = 0$. Assuming that the y coordinate is ignorable, the magnetic field $\mathbf{B}$ may be written as

$$\mathbf{B} = \nabla\psi \times \hat{y}, \qquad (3)$$

where ψ is a flux function. In equilibrium, for a charged particle of type α, the energy H_α and the momentum $P_{\alpha y}$ are conserved. Assuming that there is no equilibrium electrostatic field, these constants are

$$H_\alpha = \frac{1}{2} m_\alpha v_\alpha^2, \qquad (4)$$

$$P_{\alpha y} = m_\alpha v_{\alpha y} + q_\alpha \psi/c, \qquad (5)$$

where m_α is the mass of the charge particle of type α, q_α, its charge, and c is the speed of light. An equilibrium distribution function can be written as

$$f_\alpha = f_\alpha(H_\alpha, P_{\alpha y}),$$

$$= n_0\left(\frac{m_\alpha}{2\pi T_\alpha}\right)^{3/2} \exp\left(\frac{P_{\alpha y} V_{\alpha y} - H_\alpha}{T_\alpha}\right), \qquad (6)$$

where V_α is the drift velocity, T_α is the temperature (in energy units) for particles of type α ($= e, i$ for a hydrogen plasma), and n_0 is the average number density of electrons and protons. The temperature T_α is taken to be a constant. From the relation

$$n_\alpha = \int dv_\alpha f_\alpha, \qquad (7)$$

and the requirement of quasi-neutrality, $n_e = n_i$, we obtain the condition $V_{iy}/T_i = -V_{ey}/T_e$. The y component of Ampere's law gives

$$\nabla^2\psi = -\frac{4\pi}{c}\sum_\alpha q_\alpha \int dv_\alpha v_y f_\alpha ,$$

$$= -\frac{B_0}{\lambda}\exp\left\{\frac{2\psi}{B_0\lambda}\right\}, \quad (8)$$

where

$$\lambda \equiv \frac{2c\,(T_i+T_e)}{eB_0\,(V_{iy}-V_{ey})}, \quad (9)$$

defines a characteristic equilibrium length scale, and the constant B_0 is determined by the relation

$$B_0^2/8\pi = n_0\,(T_e+T_i) . \quad (10)$$

B_0 is a measure of the lobe magnetic field. We scale the variables $x/\lambda \to x$, $\psi/B_0\lambda \to \psi$, $p/(B_0^2/4\pi) \to p$, $\mathbf{B}/B_0 \to \mathbf{B}$, and introduce a small, positive parameter $\varepsilon \sim \partial/\partial x << \partial/\partial z \sim 1$. In this approximation, a large class of equilibrium solutions of (8) is given by [*Birn et al.*, 1975; *Birn*, 1979; *Zwingmann and Schindler*, 1980; *Lembege and Pellat*, 1982; *Zwingmann*, 1983]

$$\psi \simeq -\log\{\cosh[z f(\varepsilon x)]\,/\,f(\varepsilon x)\}, \quad (11)$$

where $f(\varepsilon x)$ is a slowly varying function of x. It follows that

$$B_x = -\frac{\partial\psi}{\partial z} = f(\varepsilon x)\tanh[z f(\varepsilon x)] , \quad (12)$$

$$B_z \equiv B_n = \frac{\partial\psi}{\partial x} = \varepsilon\left\{\frac{f'(\varepsilon x)}{f(\varepsilon x)} - z f'(\varepsilon x)\tanh[z f(\varepsilon x)]\right\}, \quad (13)$$

where prime denotes differentiation with respect to the argument. Following *Lembege and Pellat* [1982], we take $f(\varepsilon x) = \exp(\varepsilon x)$ $(x < 0)$. As a first approximation, almost all analytical studies replace (12) and (13) with

$$B_x = \tanh z , \quad (14)$$

$$B_n = \varepsilon . \quad (15)$$

However, this is valid only in the region $|x| \lesssim 1$, $|z| \to 0$. It cannot be assumed that this approximation holds for large x close to the $z = 0$ line, for the equilibrium pressure balance condition implies that $\partial p/\partial x = -\varepsilon$ which, in turn, gives $p = p_0 - \varepsilon x$, with p_0 constant. This means that the pressure increases with x, which is unrealistic for the distant tail. A better approximation for B_n than (15) is

$$B_n = \varepsilon\,(1 - z\tanh z) . \quad (16)$$

However, (16) does not represent realistically the x dependence of $B_n(x,z)$ which should tend to zero as $|x| \to \infty$. It is possible to improve on (16) by taking

$$f(\varepsilon x) = \exp[\varepsilon x/(1-\varepsilon x)] \quad (x < 0) . \quad (17)$$

We recall that the distance x is measured in units of $\lambda \sim 1\,R_E$. For specificity, we take $\varepsilon = 0.1$, and define $L_0 \equiv \varepsilon^{-1} = 10$. The scale L_0 ($\sim 10R_E$) enables us to define three separate regions:

1. Near-Earth region, $|x| << L_0$

In this region, using (17) in (12) and (13), we get

$$B_x \simeq \tanh z , \quad (18)$$

$$B_n \simeq \varepsilon\,(1 - z\tanh z) . \quad (19)$$

2. Middle region, $|x| \sim L_0$

In this region, since $\varepsilon x \sim 1$, we get

$$B_x \simeq e^{1/2}\tanh(e^{1/2} z) , \quad (20)$$

$$B_n \simeq \frac{\varepsilon}{4}\,[1 - e^{1/2} z\tanh(e^{1/2} z)] . \quad (21)$$

If we define $\mathbf{B}' \equiv e^{-1/2}\mathbf{B}$, $z' \equiv e^{1/2}z$ and $\varepsilon' \equiv e^{-1/2}\,\varepsilon/4$, then (20) and (21) become

$$B_x' \simeq \tanh z', \quad (22)$$

$$B_n' \simeq \varepsilon'\,(1 - z'\tanh z') , \quad (23)$$

which is the same as (18) and (19). In other words, the middle region has the same structure as the near-Earth region, except that the magnetic field in it is somewhat weaker and the current sheet is somewhat wider.

3. Distant-tail region, $|x| >> L_0$

In the distant-tail region, taking the limit $x/L_0 \to -\infty$, we get

$$B_x \simeq e^{-1}\tanh(e^{-1} z) , \quad (24)$$

$$B_n \simeq 0 . \quad (25)$$

Equations (24) and (25) describe essentially the configuration first considered by *Coppi et al.* [1966], who found electron tearing modes to be unstable. We caution that though the qualitative features of the magnetic field as described by (24) and (25) are reasonable, the far distant-tail region is outside the strict domain of validity of the asymptotic solution (13). Here we do not pursue this matter further, for the main focus of this paper is on the investigation of collisionless tearing modes in the near-Earth and middle regions. We also note that the near-Earth configuration described above is not totally realistic for $|x| \lesssim 5R_E$ because at these distances, the Earth's dipole field, not included in the model, plays a dominant role. Matching to the dipole field can be carried out, in principle, but is not germane to our considerations here. Figure 1 shows a plot of our asymptotic two-dimensional model with the origin set (arbitrarily) at $x = -5\,R_E$, $z = 0$.

We now discuss some features of the single-particle orbits. For a particle of mass m_α and charge q_α gyrating in magnetic field B, the Larmor frequency is $\omega_{c\alpha} = q_\alpha B/m_\alpha c$, and the typical Larmor radius is $\rho_{c\alpha} = v_{t\alpha}/\omega_{c\alpha}$, where $v_{t\alpha} \equiv (2T_\alpha/m_\alpha)^{1/2}$ is the thermal velocity. Using typical tail parameters (see, for instance, *Lui* [1987]), we get $\rho_{ci} \sim (0.5 - 1)R_E$. In both regions 1 and 2, $\rho_{ci} \geq 1$, $\rho_{ce} << 1$ (scaled by λ). Hence electrons may be treated in the guiding-center approximation, but the ions are essentially unmagnetized. Because of the z dependence of $B_n(x,z)$, the field lines are more stretched as z increases than in the case of constant B_n (see Figure 1.). In the Appendix, we show that this has the consequence that the average bounce period of electrons in both regions 2 and 3 can be substantially larger than the bounce period with constant B_n.

It is interesting to note that the large separation in the magnitude of the Larmor frequency ω_{ce} and the average bounce frequency $\omega_{be} \equiv 2\pi\tau_{be}^{-1}$ in our asymptotic equilibria (in those regions where B_n is weak) diminishes the possibility of low-

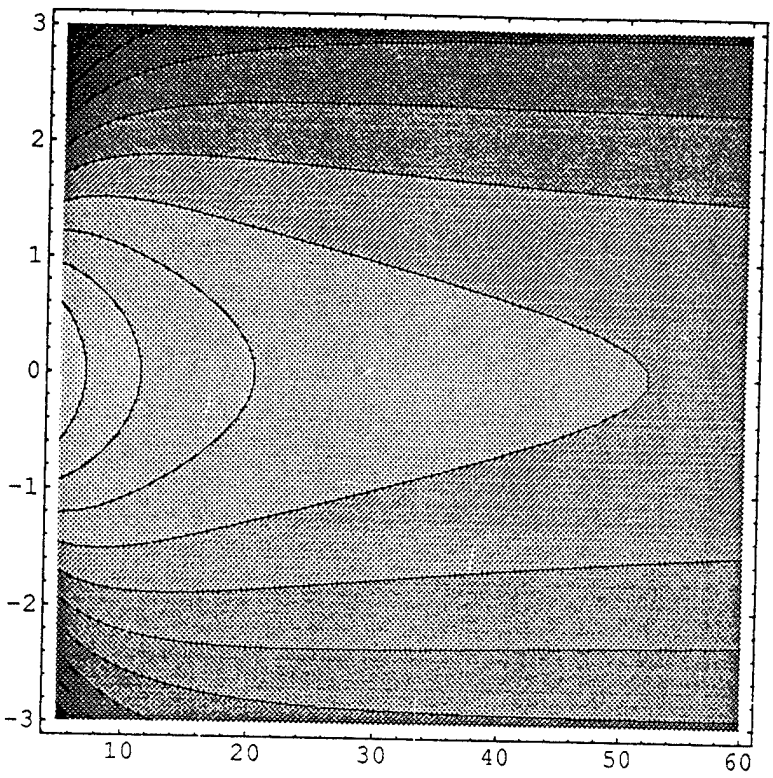

Fig. 1. Magnetic field lines in the x-z plane for equilibria with $B_n = B_n(x,z)$. The abscissae indicates $-x$, the ordinates, z.

order resonances between the Larmor and bounce frequencies for most particles, and the chaos that can result from such resonances [*Chen and Palmadesso*, 1986; *Büchner and Zelenyi*, 1987]. Hence, for equilibria with spatially varying B_n, we will not concern ourselves here with intrinsic stochastic diffusion as a possible mechanism for the restoration of ion tearing. As to whether stochastic diffusion can destabilize the ion tearing if B_n is constant has been the subject of debate recently [*Pellat et al.*, 1991; *Kuznetsova and Zelenyi*, 1991] and is an issue we shall address in section 4.

2.2. *Three-Component Equilibria*

We now consider equilibria which are symmetric in y but with all three components of the magnetic field nonzero. The magnetic field $\mathbf{B}$ is represented as

$$\mathbf{B} = B_y \hat{y} + \nabla\psi \times \hat{y} . \tag{26}$$

The condition of magnetostatic equilibrium gives the Grad-Shafranov equation [*Voigt and Hilmer*, 1987; *Paranicas and Bhattacharjee*, 1989; *Hau and Voigt*, 1992]

$$\nabla^2\psi + \frac{d}{d\psi}\left(p + \frac{B_y^2}{2}\right) = 0 , \tag{27}$$

where $p = p(\psi)$ and $B_y = B_y(\psi)$ are two free functions. If we set $B_y = 0$ and take $p = \exp(2\psi)$, we recover (8) (in dimensionless variables). We note that (8) also holds for $B_y =$ const, in which case a class of asymptotic solutions can be again constructed using (11). The B_x and B_z components for this class of solutions has already been given in section 2.1.

We now consider the effect of a nonzero $dB_y^2/d\psi$ on the solution (11). Since ψ is small and negative near $z = 0$, we make the expansion

$$B_y^2 = b_0^2 - 2b_1\psi + \ldots . \tag{28}$$

when b_0 and b_1 are positive constants. If we order $b_0 \sim b_1 \sim |\psi| \sim \varepsilon$, then near $z = 0$ (11) can be modified as

$$\psi \simeq -\log\{\cosh[zf(\varepsilon x)]/f(\varepsilon x)\} + b_1 x^2/2 , \tag{29}$$

with $f(\varepsilon x)$ specified by (17). This yields $B_n \simeq \varepsilon + b_1 x$ which implies that B_n vanishes on the $z = 0$ plane at $x \simeq -b_1/\varepsilon$. Hence an X-point appears on the $z = 0$ line (in the x-z plane) at near-Earth distances ($5\ R_E \lesssim |x| \lesssim 10\ R_E$). That the spatial dependence of B_y in a magnetostatic equilibrium can lead to the reduction of B_n to zero on the $z = 0$ plane has been recognized by *Hau and Voigt* [1992]. However, for their class of equilibrium profiles, they find that the B_y required to reduce B_n to zero is much larger than the observed average B_y in the plasma sheet. For our class of pressure and B_y profiles, this limitation is overcome because the average values of both B_n and B_y are of the same order.

The particle orbits and drift motions in the presence of B_y has been considered by WBL and will not be repeated here. Simple considerations of particle orbits near $z = 0$, where (2) holds, shows that the main stabilizing effect of B_n is to remove electrons from the $z = 0$ plane where the separatrix tends to form in the absence of B_n. The addition of the B_y field provides the electrons with another guided channel for motion near the $z = 0$ plane. Then, as noted by WBL, the growth of the electron tearing instability depends on the competing effects of B_y and B_z. As far as the effect of electron bounce is concerned, we show in the Appendix that the bounce period in the stretched magnetotail increases due to the x dependence of B_n; hence the condition $\gamma\tau_{be} \gtrsim 1$ is satisfied in the middle region. This, in turn, implies that the stabilizing effect of the electron bounce is

weakened, and that the electron tearing mode can grow in the linear regime not only for $B_y >> B_z$ but also for $B_y \sim B_z$ near the $z = 0$ plane. However, we repeat for emphasis that in the latter case, the growth rate is small, and the mode is likely to saturate nonlinearly at a low amplitude.

In this section we have made use of dimensionless variables in order to keep the notation simple. In the remainder of the paper, we shall return to using the primitive physical variables.

3. Energy Integral

In WBL, the stability of the plasma sheet was analyzed by asymptotic matching of the normal mode equations in the inner region, where finite particle inertia provides the reconnection mechanism, with the equations in the outer region, governed by ideal MHD. The technical details of such an approach are somewhat different from those involved in the Lyapunov functional method (developed by *Laval et al.* [1966]) which relies on the existence of an energy integral. We review, at first, the stability criteria that follow from the energy integral. From the linearized Maxwell's equations, it follows that

$$\int dx \left[\frac{\partial}{\partial t} \left\{ \frac{1}{8\pi} \left(\mathbf{B}_1^2 + \mathbf{E}_1^2 \right) \right\} + \mathbf{J}_1 \cdot \mathbf{E}_1 \right] = - \oint d\mathbf{a} \cdot \frac{c}{4\pi} \mathbf{E}_1 \times \mathbf{B}_1 , \tag{30}$$

where all perturbed quantities are designated by the subscript 1. We assume that the boundary conditions on the surface bounding the plasma volume ensure that the surface term vanishes. Then, it can be shown [*Laval et al.*, 1966] that

$$\int dx \, \mathbf{J}_1 \cdot \mathbf{E}_1 = - \frac{\partial}{\partial t} \int dx \, \frac{1}{2} \sum_\alpha \left[\frac{q_\alpha}{c} \int dv \, v_y \frac{\partial f_\alpha}{\partial \psi} \psi_1^2 + \int dv \, \tilde{f}_{1\alpha}{}^2 \Big/ \frac{\partial f_\alpha}{\partial H_\alpha} \right], \tag{31}$$

where

$$\tilde{f}_{1\alpha} \equiv f_{1\alpha} - \frac{\partial f_\alpha}{\partial \psi} \psi_1 . \tag{32}$$

From (30) and (31), it follows that

$$\frac{\partial}{\partial t} \delta^2 \varepsilon = 0 , \tag{33}$$

where

$$\delta^2 \varepsilon \equiv \int dx \left[\frac{1}{8\pi} \left\{ \mathbf{B}_1^2 + \mathbf{E}_1^2 \right\} - \frac{1}{2} \sum_\alpha \left\{ \frac{q_\alpha}{c} \int dv v_y \frac{\partial f_\alpha}{\partial \psi} \psi_1^2 + \int dv \, \tilde{f}_{1\alpha}{}^2 \Big/ \frac{\partial f_\alpha}{\partial H_\alpha} \right\} \right] . \tag{34}$$

The energy integral $\delta^2\varepsilon$ is a quadratic form. If $\delta^2\varepsilon$ is positive-definite for all nontrivial permissible perturbations, then the system is stable [*Kruskal and Oberman*, 1958; *Laval et al.*, 1966]. In other words, a sufficient condition for stability is

$$\delta^2 \varepsilon \geq 0 . \tag{35}$$

Furthermore, for a Maxwellian distribution, since, $\partial f_\alpha / \partial H_\alpha = -f_\alpha / T_\alpha$, we get

$$- \frac{1}{2} \sum_\alpha \int dv \, \tilde{f}_{1\alpha}^2 \Big/ \frac{\partial f_\alpha}{\partial H_\alpha} = \frac{1}{2} \sum_\alpha T_\alpha \int dv \, \frac{\tilde{f}_{1\alpha}^2}{f_\alpha} \geq 0 . \tag{36}$$

Equation (36) implies that for a Maxwellian distribution a sufficient condition for stability is

$$\delta^2 W \equiv \int dx \left[\frac{1}{8\pi} \left\{ \mathbf{B}_1^2 + \mathbf{E}_1^2 \right\} - \frac{1}{2} \sum_\alpha \frac{q_\alpha}{c} \int dv \, v_y \frac{\partial f_\alpha}{\partial \psi} \psi_1^2 \right] \geq 0. \tag{37}$$

Since collisionless tearing modes have low frequency, the electrical energy $E_1^2/8\pi$ is much smaller than the magnetic energy $B_1^2/8\pi$, and can be neglected. The sufficient condition (37) can be rewritten as

$$\delta^2 W \equiv \int dx \left[\frac{\mathbf{B}_1^2}{8\pi} - \frac{1}{2} \sum_\alpha \frac{q_\alpha}{c} \int dv \, v_y \frac{\partial f_\alpha}{\partial \psi} \psi_1^2 \right] \geq 0 . \tag{38}$$

One of the difficulties presented by the energy integral $\delta^2\varepsilon$ is that the physical interpretation of some terms is not transparent. We take for instance, the last term on the right-hand-side of (34). *Lembege and Pellat* (1982) showed, by using a Schwartz inequality, that the term has a lower bound which can be attributed to the compressibility of the electron fluid. This interpretation has been invoked repeatedly in the literature, but it is worth noting that it was meant to hold for the lower bound, and not the term itself. In fact, it is clear that there is more to the original term than electron compressibility. The function $\delta^2\varepsilon$ represents the second variation of the total energy ε which is a sum of the electromagnetic field energy and the kinetic energy of the fluid. Since the perturbed kinetic energy of the fluid is always positive definite, it must be associated with a manifestly positive definite term in $\delta^2\varepsilon$. Thus it would seem that the last term in (34) should involve the perturbed kinetic energy of the fluid plasma, but this is not obvious. (The physical interpretation of the kinetic terms is much more transparent in the energy principle of *Kruskal and Oberman* [1958] for a guiding-center plasma, but this is not the model underlying the functional (34) which has been derived from the full Vlasov equation.)

In view of the interpretational difficulties of a fully kinetic treatment, we propose a different approach for the study of collisionless tearing modes in the magnetotail. This approach uses fluid equations, and the mechanism for reconnection is provided by finite particle inertia in the generalized Ohm's law. That this is a reasonable model for collisionless tearing modes was suggested by *Furth* [1962, 1964]. The point of view we adopt here is that the energetics of collisionless tearing modes is describable within a fluid model by using a generalized Ohm's law. The wave-particle resonance condition (which is a kinetic effect) is included in such a model by simply equating the growth rate γ to kv_{te}, where k is the wave number and v_{te} the electron thermal speed. (For sheared configurations, k is replaced by $k_\parallel$, the component of k parallel to the magnetic field.) This ad hoc resonance condition misses the detailed structure of the particle distribution functions. However, we shall demonstrate that the growth rate calculated from the model equations agrees, except for an overall multiplicative factor of $\pi^{1/2}$, with the results of a fully kinetic treatment. One of the advantages of the fluid model is that the energetics of the instability is much easier to interpret

physically. This will enable us to formulate a stability condition which is both necessary and sufficient.

The linearized fluid equations are

$$\rho \frac{\partial \mathbf{u}}{\partial t} = \frac{\mathbf{J}\times\mathbf{B}_1}{c} + \frac{\mathbf{J}_1\times\mathbf{B}}{c} - \nabla p_1 , \tag{39}$$

$$\frac{\partial \rho_1}{\partial t} + \nabla\cdot(\rho\mathbf{u}) = 0 , \tag{40}$$

$$\mathbf{E}_1 + \frac{\mathbf{u}\times\mathbf{B}}{c} = \frac{m_e}{ne^2}\frac{\partial \mathbf{J}_1}{\partial t} , \tag{41}$$

$$\nabla\times\mathbf{E}_1 = -\frac{1}{c}\frac{\partial \mathbf{B}_1}{\partial t} , \tag{42}$$

$$\nabla\times\mathbf{B}_1 = \frac{4\pi}{c}\mathbf{J}_1 , \tag{43}$$

$$p_1 = n_1(T_e + T_i) , \tag{44}$$

where ρ_1 is the perturbed mass density, $\mathbf{u}$ is the perturbed fluid velocity, p_1 is the perturbed pressure, n_1 is the perturbed number density ($n_1 = n_{e1} = n_{i1}$ by quasi-neutrality), and the unsubscripted variables represent equilibrium quantities. We note that the generalized Ohm's law (41) has a term proportional to the electron mass, but none proportional to the ion mass. This can be readily seen by considering the more general form (see, for instance, *Krall and Trivelpiece* [1986])

$$\frac{\partial \mathbf{J}_1}{\partial t} = ne^2\left(\frac{1}{m_e}+\frac{1}{m_i}\right)\left(\mathbf{E}_1 + \frac{\mathbf{u}\times\mathbf{B}}{c}\right),$$

which reduces in the limit $m_e/m_i << 1$ to the form (41). Apart from terms involving finite particle inertia, the generalized Ohm's law contains other terms such as the Hall term, the electron pressure gradient term, as well as terms involving the anisotropic stress tensor. It can be shown that the first two do not qualitatively change our results for the class of equilibria considered here. However, anisotropies in the stress tensor, which are an additional source of free energy, may alter some of our conclusions. We do not consider pressure anisotropies here because it is questionable whether an instability that is primarily driven by such a source of free energy should be classified as a tearing mode.

We now use (39) – (44) to calculate the different terms in (30) (neglecting, of course, the term $E_1^2/8\pi$ which is much smaller than the term $B_1^2/8\pi$). We get

$$\mathbf{J}_1\cdot\mathbf{E}_1 = \frac{\mathbf{J}_1\times\mathbf{B}\cdot\mathbf{u}}{c} + \frac{\partial}{\partial t}\left(\frac{m_e}{2ne^2}\mathbf{J}_1^2\right). \tag{45}$$

The first term on the right-hand-side of (45) can be calculated from the momentum equation (39) which gives

$$\frac{\mathbf{J}_1\times\mathbf{B}\cdot\mathbf{u}}{c} = \frac{\partial}{\partial t}\left(\frac{1}{2}\rho u^2\right) + \nabla\cdot(p_1\mathbf{u}) - p_1\nabla\cdot\mathbf{u} + \mathbf{u}\cdot\frac{\mathbf{J}\times\mathbf{B}_1}{c} . \tag{46}$$

Writing

$$\mathbf{B}_1 = \nabla\psi_1\times\hat{y} + B_{y1}\hat{y} , \tag{47}$$

we get

$$\mathbf{u}\cdot\frac{\mathbf{J}\times\mathbf{B}_1}{c} = \nabla\cdot\left(\mathbf{u}\frac{J\psi_1}{c}\right) - \frac{J\psi_1}{c}\nabla\cdot\mathbf{u} + \frac{\psi_1}{c}\mathbf{u}\cdot\nabla J . \tag{48}$$

Since $\partial\psi_1/\partial t + \mathbf{u}\cdot\nabla\psi = 0$, and $J = J(\psi)$, we obtain

$$\mathbf{u}\cdot\nabla J = -\frac{\partial\psi_1}{\partial t}J'(\psi) \simeq \frac{c}{4\pi}\frac{\partial\psi_1}{\partial t}\frac{\psi'''(z,\varepsilon x)}{\psi'(z,\varepsilon x)}, \tag{49}$$

where the prime in $\psi'(z,\varepsilon x)$ indicates differentiation with respect to z. Defining $F \equiv \psi'(z,\varepsilon x)$ and substituting (46) – (49) in (45), we get

$$\mathbf{J}_1\cdot\mathbf{E}_1 = \frac{\partial}{\partial t}\left[\frac{1}{2}\rho u^2 + 2\pi\frac{J_1^2}{\omega_p^2} + \frac{F''}{8\pi F}\psi_1^2\right] + \nabla\cdot(\tilde{p}_1\mathbf{u}) - \tilde{p}_1\nabla\cdot\mathbf{u} , \tag{50}$$

where $\omega_p^2 = 4\pi ne^2/m_e$ is the plasma frequency, and

$$\tilde{p}_1 = p_1 - p'(\psi)\psi_1 = p_1 - J\psi_1/c . \tag{51}$$

Equation (30) can then be cast in the form,

$$\frac{\partial}{\partial t}\delta^2\varepsilon + \oint d\mathbf{a}\cdot\left[\frac{c}{4\pi}\mathbf{E}_1\times\mathbf{B}_1 + \tilde{p}_1\mathbf{u}\right] = 0 , \tag{52}$$

where

$$\delta^2\varepsilon = \delta^2W_f + \delta^2W_c + \delta^2K + \delta^2Q \equiv \delta^2U + \delta^2K + \delta^2Q . \tag{53}$$

Here

$$\delta^2W_f = \frac{1}{8\pi}\int dx\left(B_1^2 + \frac{F''}{F}\psi_1^2\right), \tag{54}$$

is the free energy of the magnetic field;

$$\frac{\partial}{\partial t}\delta^2W_c \equiv -\int dx\,\tilde{p}_1\nabla\cdot\mathbf{u} , \tag{55}$$

can be attributed to plasma compression, and

$$\delta^2Q \equiv \int dx\left(2\pi/\omega_p^2\right)J_1^2 , \tag{56}$$

is the "dissipation" due to finite electron inertia. Note that the fluid kinetic energy $\delta^2K \equiv \int dx\,(1/2)\rho u^2$, and the dissipation δ^2Q are positive definite quantities. If the boundary in (52) is chosen such that the surface term vanishes, then a sufficient condition for stability is

$$\delta^2U \equiv \delta^2W_f + \delta^2W_c \geq 0 . \tag{57}$$

In order to develop confidence in the fluid model, we now benchmark it with standard results from kinetic theory. For simplicity, we consider an incompressible fluid for which $\delta^2W_c = 0$. Then, there are three terms in the energy integral (53) among which δ^2K is the kinetic energy and always positive definite. It is clear that a tearing instability can occur if and only if there is magnetic free energy available, i.e., $\delta^2W_f < 0$, and there exists simultaneously, a mechanism for dissipation causing δ^2Q (which is always positive definite).

Let us first consider the equilibrium configuration (2) with $B_n = 0$. For this case the growth rate of the collisionless tearing instability is known analytically. The mathematical problem divides itself neatly into two regions: the outer region, away from $z = 0$, where the plasma is in quasistatic equilibrium, and the inner region near $z = 0$ where inertial and dissipative effects are important. We rewrite (54) in the form

$$\delta^2 W_f = \frac{1}{16\pi} \int d\mathbf{x} \left[\tilde{\psi}'^2 + k^2 \tilde{\psi}^2 + \frac{F''}{F} \tilde{\psi}^2 \right], \qquad (58)$$

where $\psi_1 = \tilde{\psi}(z, t) \cos kx$, $F \equiv \tanh z/\lambda$ and prime denotes derivative with respect to z. In the outer region, neglecting inertia and dissipation, we get

$$\tilde{\psi}'' - k^2 \tilde{\psi} - (F''/F) \tilde{\psi} = 0. \qquad (59)$$

Furth [1963] observed that if the first term in the integrand of (58) is integrated by parts and (59) is used, $\delta^2 W_f$ reduces to

$$\delta^2 W_f = -\frac{a_{xy}}{16\pi} \tilde{\psi} \tilde{\psi}' \Big|_{z=0-}^{z=0+},$$

$$= -\frac{a_{xy}}{16\pi} \Delta_0' \tilde{\psi}^2(0), \qquad (60)$$

where $\int d\mathbf{x} = a_{xy} \int dz$, $a_{xy} = \int dx dy$ and the parameter Δ_0' is defined by the relation,

$$\Delta_0' \equiv \frac{\tilde{\psi}'(0+) - \tilde{\psi}'(0-)}{\tilde{\psi}(0)}. \qquad (61)$$

(In obtaining (60), we have used the boundary conditions $\tilde{\psi}(-\infty) = \tilde{\psi}(+\infty) = 0$.) For the equilibrium (2) with $B_n = 0$, we get (see, for instance, WBL)

$$\Delta_0' = \frac{2}{k\lambda^2} (1 - k^2\lambda^2). \qquad (62)$$

On the interface between the inner and outer regions, the Poynting flux is

$$\oint d\mathbf{a} \cdot \frac{c}{4\pi} \mathbf{E}_1 \times \mathbf{B}_1$$

$$= \frac{c}{4\pi} \left[\left\{ \int_{z=0-} dx dy - \int_{z=0+} dx dy \right\} \hat{\mathbf{z}} \cdot \mathbf{E}_1 \times \mathbf{B}_1 \right]$$

$$= \frac{a_{xy}}{16\pi} \Delta_0' \frac{\partial}{\partial t} \tilde{\psi}^2(0) = -\frac{\partial}{\partial t} \delta^2 \varepsilon_{\text{outer}}. \qquad (63)$$

Thus the magnetic free energy in the outer region is spent as dissipation and kinetic energy in the inner region.

To determine the dissipation in the inner region, we use the generalized Ohm's law,

$$-\frac{1}{c} \frac{\partial \tilde{\psi}}{\partial t} \simeq \frac{m_e}{ne^2} \frac{\partial \tilde{J}_y}{\partial t}, \qquad (64)$$

where $J_{1y} = \tilde{J}_y \cos kx$. Equation (64) gives

$$\tilde{J}_y = -\frac{c}{4\pi} k_0^2 \tilde{\psi}, \qquad (65)$$

where $k_0 \equiv \omega_p/c$. The dissipation caused by electron inertia in the inner region is

$$\delta^2 Q \simeq a_{xy} \frac{k_0^2}{16\pi} \int_{z=-d_e}^{z=+d_e} dz\, \tilde{\psi}^2$$

$$\simeq a_{xy} \frac{k_0^2}{8\pi} d_e \tilde{\psi}(0)^2, \qquad (66)$$

where d_e is the width of the tearing layer.

Requiring that the system energy be equal to its equilibrium value [*Kruskal and Oberman*, 1958], we get

$$\delta^2 \varepsilon = \delta^2 W_f + \delta^2 K + \delta^2 Q = 0. \qquad (67)$$

For the electron tearing instability, the fluid kinetic energy $\delta^2 K$, which is dominantly due to ions, is much smaller than the electron dissipation $\delta^2 Q$. Hence, from the relation

$$\delta^2 W_f + \delta^2 Q \simeq 0, \qquad (68)$$

we obtain the tearing layer width [*Drake and Lee*, 1977]

$$d_e = d_e(k) = \frac{\Delta_0'}{2k_0^2}. \qquad (69)$$

The growth rate of the instability can be determined by using the wave-particle resonance condition. In the case $B_y = 0$ this condition is

$$\gamma \simeq k v_{te}, \qquad (70)$$

where k is determined from (62) by the requirement $\Delta_0' > 0$ ($\delta^2 W_f < 0$). From (62) and (70) we obtain

$$\gamma = \frac{v_{te}}{k_0^2 \lambda^2 d_e} (1 - k^2\lambda^2), \qquad (71)$$

$$= \left(\frac{v_{te}}{\lambda}\right) \left(\frac{\rho_e}{\lambda}\right)^{3/2} \left(1 + \frac{T_i}{T_e}\right) (1 - k^2\lambda^2). \qquad (72)$$

In writing (72) we have made use of the equilibrium relation (10). The growth rate (72) is larger by the factor $\pi^{1/2}$ than the result obtained from kinetic theory by *Laval et al.* [1966].

In the case $B_y \neq 0$ the resonance condition (70) is modified to

$$\gamma \simeq k_{\parallel} v_{te}, \qquad (73)$$

where $k_{\parallel} = k d_e/\lambda$. Equations (62) and (73) then yield

$$\gamma \simeq \frac{k v_{te} \Delta_0'}{2k_0^2 \lambda}, \qquad (74)$$

which is also larger by the factor $\pi^{1/2}$ than the result of *Drake and Lee* [1977].

Equations (71) and (74) demonstrate that the fluid model is reliable as a predictor of the parametric dependencies and the order of magnitude of the growth rate of collisionless tearing modes when $B_z = 0$, both with and without B_y. The numerical factor missed by the model has to do with the precise details of the electron distribution function.

In the last paragraph, we have used the term "collisionless tearing" instead of "electron tearing," though the latter name is commonly used. It has been shown elsewhere (for the case $B_y \neq 0$) that the full ion dynamics gives small logarithmic corrections to (74) (WBL, Appendix A]. In other words, the result (74) already includes the ion response, and there is no ion tearing branch of the dispersion equation. That being so, it is redundant to call this tearing instability electron tearing because in this case there is no other form of tearing.

4. Effect of Constant B_n

So far, we have benchmarked the fluid model by reproducing known results from kinetic theory for equilibria with $B_n = 0$. We presume that a fluid model which has been demonstrated to be reliable for $B_n = 0$ will work for equilibria with $B_n \neq 0$, representative of the Earth's magnetotail. As discussed in WBL, what will change in the presence of a nonzero B_n are the particle orbits. We consider, at first, the case B_n = constant (equation (15)) which has engendered considerable controversy in the literature.

In the presence of B_n the perturbed current has a component J_{1z} in the z direction which generates a perturbed magnetic field component B_{1y}. In WBL, this component has been calculated by writing $\mathbf{B}_1 = \nabla \times \mathbf{A}_1$, using the Coulomb gauge $\nabla \cdot \mathbf{A}_1 = 0$ and the approximation $\partial/\partial x << \partial/\partial z$ for perturbed quantities. Here we denote $\psi_1 \equiv A_{1y}$, $\chi_1 \equiv A_{1z}$, and write

$$B_1^2 \equiv \psi_1'^2 + \left(\frac{\partial \psi_1}{\partial x}\right)^2 + B_{1y}^2 \tag{75}$$

Substituting (75) in (54), we have

$$\delta^2 W_f = \frac{1}{8\pi} \int dx \left[\psi_1'^2 + \left(\frac{\partial \psi_1}{\partial x}\right)^2 + \frac{F''}{F} \psi_1^2 + B_{1y}^2 \right]. \tag{76}$$

We take $\chi_1 = \tilde{\chi}_1(z,t) \cos kx$. In the outer region, we have [WBL, Appendix B]

$$\tilde{\chi}'' - k^2 \tilde{\chi} \simeq 0, \tag{77}$$

which has a solution of the form $e^{-k|z|}$. This solution has a jump continuity in its logarithmic derivative, specified by

$$\Delta_2' \equiv \frac{\tilde{\chi}'(0+) - \tilde{\chi}'(0-)}{\tilde{\chi}(0)} \simeq -2k. \tag{78}$$

In the Coulomb gauge, we can now write $B_{1y}^2 = k^{-2} (\tilde{\chi}''' - k^2 \tilde{\chi})^2$. Then, the contribution of the last term in the integrand of (76) is $-(a_{xy}/16\pi) \Delta_2' \tilde{\chi}^2(0)$. This has a clear physical interpretation: B_n reduces the magnetic free energy available to the tearing instability and hence has a stabilizing effect.

The computation of the first three terms in the integrand of (76) is somewhat more involved. One of the complications introduced by B_n is that it introduces a phase shift that, in effect, couples the cosine solution in x, i.e., $\tilde{\psi}_c(z,t) \cos kx$ with the sine solutions in x, i.e., $\tilde{\psi}_s(z,t) \sin kx$. In the outer region equation,

$$\psi' \frac{\partial}{\partial x} \nabla^2 \psi_1 = \psi''' \frac{\partial \psi_1}{\partial x} + B_n \frac{\partial}{\partial z} \nabla^2 \psi_1, \tag{79}$$

substituting

$$\psi_1 = \tilde{\psi}_c(z,t) \cos kx + \tilde{\psi}_s(z,t) \sin kx, \tag{80}$$

we obtain

$$h''' \tilde{\psi}_c = h' (\tilde{\psi}_c'' - k^2 \psi_c) - \frac{\varepsilon}{k\lambda} (\tilde{\psi}_s''' - k^2 \tilde{\psi}_s'), \tag{81}$$

$$h''' \tilde{\psi}_s = h' (\tilde{\psi}_s'' - k^2 \tilde{\psi}_s) + \frac{\varepsilon}{k\lambda} (\tilde{\psi}_c''' - k^2 \tilde{\psi}_c'), \tag{82}$$

where $h' \equiv \lambda^{-1} \tanh z/\lambda$. Following the method outlined in Appendix B of WBL, we get

$$\tilde{\psi}_c \simeq \tilde{\psi}_c^{(0)} \pm \varepsilon \lambda b h' \tilde{\psi}_s^{(0)} + 0(\varepsilon^2), \tag{83}$$

$$\tilde{\psi}_s \simeq \tilde{\psi}_s^{(0)} \mp \varepsilon \lambda b h' \tilde{\psi}_c^{(0)} + 0(\varepsilon^2), \tag{84}$$

where the upper sign in (83) and (84) corresponds to $z > 0$, the lower sign to $z < 0$, and $b \equiv 4/(k \Delta_0' \lambda^2)$. The leading order solutions $\tilde{\psi}_c^{(0)}$ and $\tilde{\psi}_s^{(0)}$ obey

$$h''' \tilde{\psi}_{c,s}^{(0)} = h' \left(\tilde{\psi}_{c,s}^{(0)''} - k^2 \lambda^2 \tilde{\psi}_{c,s}^{(0)}\right). \tag{85}$$

We now use (81) – (85) to calculate the first three terms in $\delta^2 W_f$ These are

$$\delta^2 W_\psi \equiv \frac{1}{8\pi} \int dx \left[\psi_1'^2 + \left(\frac{\partial \psi_1}{\partial x}\right)^2 + (F''/F) \psi_1^2 \right],$$

$$= \frac{a_{xy}}{16\pi} \int dx \left[\tilde{\psi}_c'^2 + k^2 \tilde{\psi}_c^2 + \frac{h'''}{h'} \tilde{\psi}_c^2 \right.$$

$$\left. + \tilde{\psi}_s'^2 + k^2 \tilde{\psi}_s^2 + \frac{h'''}{h'} \tilde{\psi}_s^2 \right]. \tag{86}$$

Integrating the right-hand side of (86) by parts and using (81) and (82), we get

$$\delta^2 W_\psi \simeq -\frac{a_{xy}}{16\pi} \Delta_0' \tilde{\psi}^2(0) + \frac{\varepsilon}{k\lambda} \left[\int dz \frac{\tilde{\psi}_c^{(0)}}{h'} \left(\frac{h'''}{h'} \tilde{\psi}_s^{(0)}\right)' \right.$$

$$\left. - \int dz \frac{\tilde{\psi}_s^{(0)}}{h'} \left(\frac{h'''}{h'} \tilde{\psi}_c^{(0)}\right)' \right], \tag{87}$$

where $\tilde{\psi}^2(0) \equiv \tilde{\psi}_c^2(0) + \psi_s^2(0)$. Since $\tilde{\psi}_c^{(0)}$ and $\tilde{\psi}_s^{(0)}$ are both solutions of (85) and obey the same boundary conditions, the last two terms of (87) cancel each other exactly. Defining

$$\Delta_r' \equiv \frac{\Delta_0' \tilde{\psi}^2(0) + \Delta_2' \tilde{\chi}^2(0)}{\tilde{\psi}^2(0) + \tilde{\chi}^2(0)}, \tag{88}$$

and using (87), we get

$$\delta^2 W_f \simeq -\frac{a_{xy}}{16\pi} \Delta_r' \left(\tilde{\psi}^2(0) + \tilde{\chi}^2(0)\right). \tag{89}$$

We now evaluate $\delta^2 W_c$, which is due to plasma compressibility. If B_n is large and constant, then the inequality $\gamma \tau_{be} << 1$ is satisfied. Under these conditions the electrons have a stabilizing compressional contribution due to the bouncing motion between mirror points along a field line. However, the effect of ion compressibility, which comes into play because of

the quasi-neutrality constraint $n_{e1} = n_{i1}$, is larger than the effect of electron compressibility. The magnetic free energy $\delta^2 W_f$ is generally not large enough to provide energy for compressing the ions, and the mode is stabilized unless the wavelength is very large. In order to demonstrate this, we calculate $\delta^2 W_c$ from the fluid equations. From the linearized continuity equation, we get

$$\nabla \cdot \mathbf{u} = -\frac{1}{n}\frac{\partial n_1}{\partial t} - \mathbf{u}\cdot\nabla\psi \frac{d}{d\psi}\ln n . \tag{90}$$

Using the equilibrium relation (7), and defining

$$\tilde{n} = n_1 - \frac{dn}{d\psi}\psi_1 , \tag{91}$$

we rewrite (90) in the form

$$\nabla\cdot\mathbf{u} = -\frac{1}{n}\frac{\partial\tilde{n}}{\partial t} - \frac{2}{B_0\lambda}\left(\frac{\partial\psi_1}{\partial t} + \mathbf{u}\cdot\nabla\psi\right). \tag{92}$$

Averaging (92) over an electron bounce period, we have

$$\langle\nabla\cdot\mathbf{u}\rangle = -\frac{\gamma}{n}\langle\tilde{n}\rangle - \frac{2}{B_0\lambda}\left(\gamma\langle\psi_1\rangle + B_n\langle u_{1x}\rangle\right), \tag{93}$$

where $\langle\ \rangle$ indicates an average over the rapid bounce motion of the electrons. Since the bounce motion involves dominantly the outer region, we neglect the electron inertia term in averaging Ohm's law (41). We then obtain

$$\langle E_{1y}\rangle = -\frac{1}{c}\left\langle\frac{\partial\psi_1}{\partial t}\right\rangle \simeq \frac{\langle v_{1x}\rangle B_n}{c} . \tag{94}$$

Using (44), (51), (93), and (94), we get

$$\langle\tilde{p}_1\ \nabla\cdot\mathbf{u}\rangle \simeq -\frac{k^2}{2B_n^2}n_0(T_e+T_i)\left\langle\frac{\partial}{\partial t}\psi_1^2\right\rangle$$

$$= \frac{k^2 B_0^2}{16\pi B_n^2}\left\langle\frac{\partial}{\partial t}\psi_1^2\right\rangle . \tag{95}$$

Hereafter, to simplify notation, we shall drop the averaging sign. Equation (95) then gives

$$\delta^2 W_c \simeq \int\frac{dx}{16\pi}\ \frac{k^2 B_0^2}{2B_n^2}\ \tilde{\psi}^2 . \tag{96}$$

Equation (96) can be rewritten as

$$\delta^2 W_c \simeq \frac{a_{xy}}{16\pi}\ \frac{k^2 B_0^2}{2B_n^2}\ z_0\tilde{\psi}^2(0) , \tag{97}$$

where

$$z_0 = 2\int_0^\infty dz\ \exp(-2kz)\left[1+\tanh z/\lambda\right]^2 . \tag{98}$$

Equation (97) agrees with (91) of *Galeev* [1984] (except for the area term a_{xy} which has been taken to be unity by Galeev without loss of generality). The constant z_0 is estimated by Galeev from physical arguments; here, we evaluate (98) asymptotically to obtain

$$z_0 \simeq 1/k . \tag{99}$$

Note that an upper bound for $\delta^2 W_f$ is obtained by setting $\tilde{\chi}(0)$ equal to zero. Hence, from (57), a sufficient condition for stability is

$$-\Delta_0' + k/\varepsilon^2 \geq 0 , \tag{100}$$

which for $\varepsilon \ll 1$, reduces to

$$k\lambda/\varepsilon \geq 2 . \tag{101}$$

Equation (101) is close to the sufficient condition for stability $k\lambda/\varepsilon \geq 4/\pi$, derived by *Lembege and Pellat* [1982]. The inequality (101) implies that all wavelengths smaller than $\pi\lambda/\varepsilon$ are stable. For example, if we take $\lambda \sim 1\ R_E$, $\varepsilon \sim 0.1$, we find that wavelengths smaller than $30\ R_E$ are stable. Of course, this does not necessarily mean that wavelengths larger than $\pi\lambda/\varepsilon$ are unstable because violation of (101) does not imply instability. If one proceeds with the hypothesis that instability is possible for $k\lambda/\varepsilon < 2$, it can be shown, following *Lembege and Pellat* [1982], that a long-wavelength ion tearing mode is impossible. We refer the reader to the work of *Lembege and Pellat* [1982] for further details.

Attempts have been made to restore the ion tearing instability by invoking pitch-angle diffusion [*Coroniti*, 1980; *Galeev*, 1984 and references therein] or intrinsic chaotic diffusion [*Büchner et al.*, 1987]. We now demonstrate that even in the presence of these effects, the most that we can get is some form of weak electron tearing and that there is no ion tearing. At first we note that occasionally, a source of some confusion in the literature has been the misleading premise that it is electron compressibility that stabilizes tearing in the presence of a constant B_n. From this premise follows the argument that if the electrons are removed by pitch angle scattering or intrinsic stochastic diffusion, then it is possible to neglect the electrons while the ions tear field-lines. Our fluid model clearly indicates that electron compressibility is less of a factor than ion compressibility for conditions typical of the magnetotail. Inspection of (95) shows that both ions and electrons contribute to $\delta^2 W_c$, but the electron contribution to $\delta^2 W_c$ may be apportioned as $[T_e/(T_i+T_e)]\ \delta^2 W_c$, whereas the ion contribution is $[T_i/(T_i+T_e)]\ \delta^2 W_c$. Since $T_i \approx 5T_e$ is typical in the magnetotail, this apportionment indicates that the dominant contribution to fluid compressibility comes from ions.

In order to pinpoint the differences between our results and others in the literature, we refer the reader to the review by *Galeev* [1984]. Galeev's equation (91) gives the energy spent for plasma compression, in agreement with our $\delta^2 W_c$. Subsequently, in the presence of pitch angle scattering, Galeev attributes a compressional term similar to $\delta^2 W_c$ entirely to electrons (see his equation (96)), and yet another contribution due to ions (his equation (97)). Our fluid model yields the compressional energy,

$$\delta^2 W_c \simeq \frac{a_{xy}}{16\pi}\ \frac{k^2 B_0^2}{2B_n^2}\ \frac{\gamma}{\gamma+\nu_{\text{eff}}}\ z_0\tilde{\psi}^2(0) , \tag{102}$$

where ν_{eff} is the bounce-averaged effective collision frequency. Note that when $\nu_{\text{eff}} = 0$, (102) reduces to (97), as it should. (If $\nu_{\text{eff}} >> \gamma$, then the factor $\gamma/(\gamma+\nu_{\text{eff}})$ can be approximated by γ/ν_{eff}). Thus $\delta^2 W_c$ (our equation (102)) includes the compressional effect due to both electrons and ions, and that there is no separate ion contribution as Galeev's equation (97) suggests. For large ν_{eff} (or equivalently, large stochastic diffusion), the stabilizing effect of $\delta^2 W_c$ can be strongly

reduced. Under these conditions, it is possible, in principle, to recover an electron tearing instability, but there is no ion tearing. This conclusion supports the recent results of *Pellat et al.* [1991] who question the very existence of ion tearing, but contradicts the findings of *Kuznetsova and Zelenyi* [1991].

In view of the controversy in analytical theories, much can be learned from particle simulations. Unfortunately, electromagnetic particle simulations of collisionless tearing inevitably involve making compromising choices on such parameters as m_e/m_i, the system size (which determines the range of unstable wavelengths), and the spatial grid size. We have cited several such simulations earlier, and it is fair to say that in all of them, an instability with the theoretically predicted growth rate and characteristics of ion tearing has been very difficult to find. Since we believe that both electron and ion dynamics (which are tied by the constraint of quasi-neutrality) should be retained in simulations of collisionless tearing, we first comment on reported results from two-species simulations that include a B_n field. *Swift and Allen* [1987, P.10,015] report that their previous unpublished work showed "no evidence of the development of any type of instability." They also attribute correctly the observed stability to ion compressibility. *Zwingmann et al.* [1990] report results mostly for the mass ratio $m_e/m_i = 1$, with some discussion of a case with $m_e/m_i = 1/10$. As they note, the case $m_e/m_i = 1$ cannot distinguish between electron and ion tearing. (If an ion tearing mode exists, its growth rate should be much larger than the electron tearing growth rate when the mass ratio is realistic.) Their results show significant discrepancies with theoretical predictions [*Schindler*, 1974]. In particular, the growth rate observed in the simulation is up to an order of magnitude less than predicted by theory. We attribute the growth of the instability in these simulations for small values of B_n to electron tearing, not ion tearing. This hypothesis can be tested, of course, by a study which computes the growth rate as a function of m_e/m_i.

Apart from two-species simulations, there are one-species simulations of the ion tearing mode in which the electrons are involved only as a static charge-neutralizing background [*Terasawa*, 1981; *Hamilton and Eastwood*, 1982; *Swift*, 1983; *Ambrosiano and Lee*, 1983; *Pritchett et al.*, 1991]. It is clear from our previous discussion that these simulations cannot realistically simulate electron tearing modes. Furthermore, any inference on the viability of the ion tearing mode from these simulations is questionable because electron dynamics has been eliminated arbitrarily for reasons of computational convenience.

We conclude this section with the remark that unless the B_n field is very small, field lines cannot reconnect to form islands in the linear regime. The striking contrast between configurations with $B_n = 0$ which tear easily to form magnetic islands and configurations in which significant values of B_n inhibit tearing is illustrated well by Figures 6.2.4 and 6.2.9 in Galeev's review paper. In Figure 6.2.4, islands develop at the separatrix where collisionless reconnection provides accessibility to a state of lower energy. In Figure 6.1.9, there is no well-defined separatrix, and the system sustains global compressional oscillations.

5. Effects of B_y and Spatially Varying B_n

WBL considered the effect of a constant B_y field superimposed on the two-dimensional configuration of section 4. Their analysis of the electron tearing mode dealt with the inner region dynamics using kinetic theory, but global aspects of the dynamics such as the bounce motion was neglected. The aim of the present effort is to explore the consequences of these global effects in the context of the improved asymptotic equilibria developed in section 2.

At first, we consider equilibria in which B_y is constant, and the spatial dependencies of B_x and B_n are described in region 1 (near-Earth) by (18) and (19), in region 2 (middle) by (20) and (21), and in region 3 (distant-tail) by (22) and (23), respectively. As noted in section 2, the spatial structure of B_n in the near-Earth and middle regions are similar, except that the magnetic field is weaker and the current sheet is wider in the middle region. It is shown in Appendix A that the bounce period τ_{be} in region 2 is larger by an order of magnitude than τ_{be} in region 1. In region 3, since B_n is vanishingly small, τ_{be} is extremely large.

Certain conditions must be fulfilled for the electron tearing instability to occur. First, there must be magnetic free energy available to drive the instability; i.e., we must have $\delta^2 W_f < 0$, where $\delta^2 W_f$ is given by (89). This means that the stability parameter Δ_r' must be positive. In standard analyses of collisionless stability of the tail (see, for instance, *Galeev* [1984, and references therein]), Δ_r' is replaced by Δ_0'. Note that this overestimates the range of unstable wavelengths because Δ_2' (equation (78)) is negative.

Second, the stabilizing compressional energy $\delta^2 W_c$ due to the bounce motion of electrons should not exceed the destabilizing term $\delta^2 W_f$. We show, a posteriori, that in region 1, $\gamma\tau_{be} << 1$, but in regions 2 and 3, we have $\gamma\tau_{be} \gtrsim 1$ and $\gamma\tau_{be} >> 1$, respectively. The compressional stabilization is thus significant in region 1 but not so in regions 2 and 3.

In regions 2 and 3, the effect of the electron bounce can be neglected. We are then back in the framework of WBL who obtained the dispersion equation for electron tearing modes neglecting electron bounce. From a kinetic analysis, carried out in Appendix B, we obtain the complex frequency $\omega = \omega_R + i\gamma$, where

$$\omega_R \simeq \omega_* - \frac{k v_{te} s_0}{2\sqrt{\pi}\, k_0^2 L_s} \Delta_i', \tag{103}$$

$$\gamma \simeq \frac{1}{\sqrt{\pi}} \frac{k v_{te} s_0}{2 k_0^2 L_s} \Delta_r', \tag{104}$$

in region 2. Here $\omega_* = k V_{ex}$, s_0 is a constant which can be shown to have a numerical value of approximately 10 under typical conditions, $L_s = \lambda(B_y^2 + B_n^2)^{1/2}/B_0$ and $\Delta_i' = 8\varepsilon/(k\Delta_0' \lambda^3) << \Delta_r'$. For comparison, we recall that if $B_n \simeq 0$, the real and imaginary parts of the complex growth rate are, respectively [*Drake and Lee*, 1977], given by

$$\omega_R \simeq \omega_*, \tag{105}$$

$$\gamma \simeq \frac{1}{\sqrt{\pi}} \frac{k v_{te}}{2 k_0^2 L_s} \Delta_0'. \tag{106}$$

In (103) and (104), all equilibrium quantities are evaluated at $z = 0$. In particular, since $\tilde{\psi}(0)/\tilde{\chi}(0) \simeq B_y/B_n(z=0)$, we can write

$$\Delta_r' \simeq \frac{\alpha^2 \Delta_0' + \varepsilon^2 \Delta_2'}{\alpha^2 + \varepsilon^2}, \tag{107}$$

where $\alpha \equiv B_y/B_0$. If $\alpha \sim \varepsilon$, a necessary condition for instability is

$$\Delta_r' = \Delta_0' + \Delta_2' > 0, \tag{108}$$

which gives

$$k\lambda < 1/\sqrt{2}. \tag{109}$$

Equation (109) implies that wavelengths larger than $2\sqrt{2}\,\pi\lambda$ ($\simeq 9\,R_E$ for $\lambda \simeq 1\,R_E$) may be unstable. In order for the instability to grow, however, it must also satisfy the condition $\gamma\tau_{be} \gtrsim 1$. A viable class of instabilities is obtained for $k^{-1} \geq 3\lambda$; these do obey the condition $\gamma\tau_{be} \gtrsim 1$ for $s_0 \simeq 10$, $\nu_{te} \simeq 2\lambda/s$.

If $B_y << B_z$, i.e., $\varepsilon >> \alpha$, the condition $\Delta_r' > 0$ for instability reduces to $k\lambda < \alpha/\varepsilon$. As shown by WBL, this condition predicts unstable wavelengths which are much too large to account for reconnection events in the near-Earth and middle regions.

We note that (104) and (106) have been obtained from a kinetic analysis, and except for a factor of $\pi^{1/2}$, can also be obtained from the fluid model. The fluid analog of (106) is (74), derived in section 3. The fluid analog of (104) has been derived in Appendix B, and is given by

$$\gamma \simeq \frac{k v_{te} s_0}{2k_0^2 L_s}\,\Delta_r'. \tag{110}$$

As before, the results from the fluid and kinetic calculations differ by a multiplicative factor of $\pi^{1/2}$.

In regions where $\gamma\tau_{be} \gtrsim 1$, the destabilizing effect of B_y may be understood as follows. If $B_y = 0$, *Galeev* [1984] points out that the energy spent for plasma compression is the work done by the perturbed plasma current generated by the perturbed pressure gradient,

$$\frac{\partial \tilde{p}_1}{\partial x} = \frac{\tilde{J}_{1y} B_n}{c}. \tag{111}$$

If $B_y \neq 0$, (111) changes to

$$\frac{\partial \tilde{p}_1}{\partial x} = \frac{1}{c}\left(\tilde{J}_{1y} B_n - \tilde{J}_{1z} B_y\right). \tag{112}$$

Now, if $\gamma\tau_{be} << 1$, then the bounce average of (112) should be taken, and the second term on the right-hand side of (112) averages to zero. On the other hand, if $\gamma\tau_{be} \gtrsim 1$, or $\gamma\tau_{be} >> 1$, then there is no bounce average to be taken, and the second term tends to reduce the first term because $\tilde{J}_{1y}/\tilde{J}_{1z} \sim B_y/B_z$.

There is some evidence in the numerical simulations of *Swift and Allen* [1987] that the presence of B_y enhances the tearing activity near $z = 0$ compared with the case $B_y = 0$. (See their section 4.3.) Clearly, there is a need for two-species simulations including B_y using either asymptotic equilibria of the kind developed in this paper, or numerical solutions of the equilibrium Grad-Shafranov equation [cf. *Voigt and Hilmer*, 1987].

Finally, we comment on equilibria with spatially varying B_y, discussed in section 2.2. There we show that, for a class of pressure profiles, the spatial variation of B_y can cause the formation of an X point on the $z = 0$ line. We show, furthermore, that this can occur at near-Earth distances for average values of B_y and B_n characteristic of the plasma sheet. The configuration thus formed is likely to be highly magnetically stressed. Under such conditions, rapid reconnection may occur at the separatrix in both the linear and nonlinear regimes. We conjecture that the collisionless reconnection rate in this geometry is likely to be much larger than the rates derived in this paper. Such a geometry calls for a separate treatment, and the exploration of that possibility is left to future work.

6. Conclusions

This paper makes two main contributions to the problem of collisionless tearing modes in the Earth's magnetotail. The first involves the development of asymptotic magnetotail equilibria including all three components of the magnetic field, with realism in the modeling of the normal component of the magnetic field, $B_n(x,z)$. The second involves the development of a fluid model that is physically transparent and accurate in reproducing the parametric dependencies of the growth rates of collisionless tearing modes calculated from kinetic theory.

One of the significant conclusions of this paper is that the ion tearing mode, which has been the subject of considerable research and controversy over the last two decades, does not occur. This is true for both two- and three-component models of the magnetotail. We are not the first to suggest this, because *Lembege and Pellat* [1982] and *Pellat et al.* [1991] have preceded us, albeit in the context of the simple two-component equilibrium (1). We find that for the two and three-component equilibria given in this paper that if there is a collisionless tearing instability in the magnetotail, it is the electron tearing mode.

There are certain conditions that must be fulfilled for the electron tearing mode to be seen. A significant value of B_n in the two-component magnetotail, represented by (1), tends to suppress the instability. The reason for this strong stabilization can be understood in dynamical terms. In our view, the dynamics are a symptom of a deeper cause which has to do with geometry. The main cause of the stabilization of the tearing mode is that B_n destroys the separatrix at $z = 0$. By contrast, if we set $B_n = 0$ but include a B_y field, the separatrix at $z = 0$ is undisturbed, and electron tearing modes can easily occur.

The main difficulty posed by the magnetotail is that all three components of the magnetic field can be significant. Then, the circumstances that favor electron tearing are those that minimize global features of the dynamics such as electron bounce, and keep the electron confined near $z = 0$. It is intuitively clear that the B_y field tends to confine electrons near $z = 0$, and hence helps the electron tearing instability grow. The asymptotic equilibria presented in this paper have regions where the stabilizing effect of electron bounce can be neglected and where the presence of $B_y \sim B_n$ can cause the excitation of electron tearing. A tearing instability in which field lines actually undergo genuine topological change does not occur unless B_n is very small. Unless topological change occurs, the instability is likely to saturate nonlinearly at a relatively low amplitude. We do not believe that such a weak instability can account for the dramatic signatures associated with current disruption and diversion during substorms.

The instability is more interesting when B_n is zero. We have demonstrated that if we include B_y and allow it to vary spatially in a three-dimensional magnetotail equilibrium, then B_n can vanish at near-Earth distances. The linear as well as the nonlinear growth of electron tearing modes in such a configuration is likely to lead to interesting results and will be investigated in the near future.

An important challenge for a theory of substorms is that it should account not only for the violent activity that is associated with substorms, but also identify conditions under which the magnetotail is stable. A universal instability that occurs always

and spontaneously is likely not to be a correct explanation because that would suggest the magnetotail is always unstable, which is not observed to be the case. In this work, we have identified conditions under which electron tearing modes may be unstable and delineated regimes when they are not. It is our hope that this paper, as well as its forerunner (WBL), will stimulate a reexamination of old as well as new data in substorms with a renewed emphasis on the B_y field. Observationalists, many of whom we have cited here, have been aware over the last 15 years of the ubiquitous presence of the B_y field, varying spatially as well as in time before and during different phases of a substorm. What is required is a more systematic study correlating B_y and B_z with the occurrence of substorm onset. More two-species electromagnetic particle simulations, including all three components of the magnetic field, are also required, both to check analytic theory and to model realistically global features of magnetotail equilibria.

APPENDIX A: BOUNCE MOTION OF ELECTRONS

An electron gyrating along a field line in the magnetotail may, under certain conditions, bounce between two turning points. Since $B_n(x, z) = B_n(x, -z)$, the z coordinates of the turning points may be written as $z = \pm z_t$, where the constant z_t is determined by the parallel energy of the electron,

$$\frac{1}{2} m v_{||}^2 = \frac{1}{2} m v^2 - \mu B \equiv K - \mu B , \tag{A1}$$

where μ is the magnetic moment. The bounce period is defined as

$$\tau_b = \oint \frac{dl}{v_{||}} , \tag{A2}$$

where l is the coordinate along a field line. Using the relation $dl/B = dl/B_z$, we have

$$\tau_b = 4 \int_0^{z_t} \left(\frac{m}{2K}\right)^{1/2} \frac{B}{B_z} \frac{dz}{(1 - \mu B/K)^{1/2}} . \tag{A3}$$

We recall that $B_x \simeq \tanh z$ and $B_n \simeq \varepsilon(1 - z \tanh z)$. For $|z| \lesssim 1$, we use the approximations $B_x \simeq z$ and $B_n \simeq \varepsilon(1 - z^2)$. Since $B_n(z = \pm 1) \simeq 0$, we need consider only the domain $0 \le |z| < 1$. Equation (A3) may be then approximated as

$$\tau_b \simeq \frac{4}{\varepsilon v_{te}} \int_0^{z_t} dz \frac{[z^2 + \varepsilon^2 (1 - z^2)^2]^{1/2}}{(1 - z^2)[1 - B(z)/B(z_t)]^{1/2}} . \tag{A4}$$

Defining $u^2 = 1 - z/z_t$, we get

$$\tau_b \simeq \frac{8 z_t^2}{\varepsilon v_{te}} \int_0^1 du \frac{1 - u^2}{1 - z_t^2 (1 - u^2)^2} . \tag{A5}$$

If we take $B_n = \varepsilon$ everywhere, as in *Lembege and Pellat* [1982], we get

$$\tau_b \simeq \frac{8 z_t^2}{\varepsilon v_{te}} \int_0^1 (1 - u^2)\, du = \frac{16}{3\varepsilon v_{te}} z_t^2 . \tag{A6}$$

For $z_t \sim 1$, equation (A6) predicts a bounce period of $\tau_b \simeq 25\ s$ (for $\varepsilon \sim 0.1$ and $v_{te} \sim 2/s$). We show below that this is much smaller than the bounce period in our asymptotic equilibria.

The integral (A5) can be evaluated exactly to give

$$\tau_b \simeq \frac{4}{\varepsilon v_{te}} \left[(1 - z_t^2)^{-1/2} \tan^{-1}\left\{ z_t / (1 - z_t^2)^{-1/2} \right\} - \frac{1}{2} (1 + z_t^2)^{-1/2} \log \frac{(1 + z_t^2)^{1/2} + z_t}{(1 + z_t^2)^{1/2} - z_t} \right] . \tag{A7}$$

Clearly, $\tau_b \to \infty$, as $|z_t| \to 1$; this means that some electrons are lost and some have very large bounce periods. An average bounce period for the confined electrons can be obtained by averaging over the distribution function of electrons. For fixed z we have

$$\langle \tau_b \rangle \approx \int_{-v_L}^{v_L} dv_{||} \left(\frac{m_e}{2\pi T_e}\right)^{1/2} \exp\left(\frac{-m v_{||}^2}{2T_e}\right) \tau_b , \tag{A8}$$

where v_L is the maximum parallel velocity above which the electron is lost. Note that (A8) underestimates the bounce period because it should be normalized by the fraction of confined electrons which is smaller than 1. Since all bouncing electrons pass through $z = 0$, we take the distribution at $z = 0$ to be Maxwellian. We estimate $v_L \approx v_{te}[B(z = 1)/\varepsilon - 1]^{1/2}$. Asymptotic evaluation of the integral (A8) gives $\langle \tau_b \rangle \approx 3e\pi/\varepsilon v_{te}$. For $\varepsilon \simeq 0.1$, $v_{te} \simeq 2/s$, we get $\langle \tau_b \rangle \approx 10^2 s$ in region 1. In region 2, since B_n is space-dependent, ε is replaced by $\varepsilon' \equiv e^{-1/2}\varepsilon/4$, and the length scale is amplified by $e^{1/2}$ (see section (2.1)). Then $\langle \tau_b \rangle$ is amplified by a numerical factor of approximately $4e^{1/2}$, which gives $\langle \tau_b \rangle \gtrsim 10^3 s$. For $\gamma \simeq 10^{-3} s$, we thus have $\gamma \langle \tau_b \rangle \ll 1$ in region 1, but $\gamma \langle \tau_b \rangle \gtrsim 1$ in region 2.

APPENDIX B: DISPERSION RELATION FOR ELECTRON TEARING MODES

In order to keep this paper self-contained, we review here the derivation of the dispersion relation for electron tearing modes using kinetic theory. The main effort lies in calculating the perturbed current $\mathbf{J}_1$ from the perturbed distribution function $f_{1\alpha}$ by means of the relation

$$\mathbf{J}_1 = \sum_\alpha q_\alpha \int d\mathbf{v}\, \mathbf{v} f_{1\alpha} . \tag{B1}$$

It is convenient to use (32) to write

$$f_{1\alpha} = \frac{\partial f_\alpha}{\partial \psi} \psi_1 + \tilde{f}_{1\alpha} . \tag{B2}$$

We recall that [*Laval et al.*, 1966; *Galeev*, 1984]

$$\sum_\alpha q_\alpha \int d\mathbf{v}\, \mathbf{v} \frac{\partial f_\alpha}{\partial \psi} \psi_1 = \frac{c}{2\pi} \frac{1}{\lambda^2 \cosh^2(z/\lambda)} \psi_1 \hat{y} . \tag{B3}$$

Hence, from Ampere's law,

$$\mathbf{J}_1 = -\frac{c}{4\pi} \nabla^2 \mathbf{A}_1 = \sum_\alpha q_\alpha \int d\mathbf{v}\, \mathbf{v} f_{1\alpha} , \tag{B4}$$

we get

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2}\right) \mathbf{A}_1 + \frac{2}{\lambda^2 \cosh^2(z/\lambda)} \psi_1 \hat{y} = -\frac{4\pi}{c} \sum_\alpha q_\alpha \int d\mathbf{v}\, \mathbf{v} \tilde{f}_{1\alpha} . \tag{B5}$$

From the linearized Vlasov equation it follows that

$$\frac{d}{dt}\tilde{f}_{1\alpha} = -\frac{q_\alpha}{T_\alpha}\,\mathbf{E}_1\cdot\mathbf{v}f_{0\alpha}, \tag{B6}$$

where $d/dt \equiv \partial/\partial t + \mathbf{v}\cdot\nabla$ and $f_{0\alpha}$ is the Maxwellian distribution. Integrating (B6) along characteristics and substituting the result in (B5), we get

$$\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial z^2}\right)A_1 + \frac{2}{\lambda^2\cosh^2(z/\lambda)}\,\psi_1\hat{y}$$

$$= -\frac{4\pi}{c}\sum_\alpha \frac{q_\alpha^2}{T_\alpha}\int d\mathbf{v}\,\mathbf{v}(t)\int_{-\infty}^{t}\mathbf{v}(t')\cdot\mathbf{E}_1(t')f_{0\alpha}(t')\,dt'$$

$$\equiv -\frac{4\pi}{c}\,\underset{\leftrightarrow}{\sigma}\cdot\mathbf{E}_1, \tag{B7}$$

where $\underset{\leftrightarrow}{\sigma}$ is the collisionless conductivity tensor which can also be written as (see, for instance, *Horton and Tajma* [1990]),

$$\underset{\leftrightarrow}{\sigma} \equiv -\sum_\alpha \frac{q_\alpha^2}{T_\alpha}\int d\mathbf{v}\int_0^\infty d\tau\,\mathbf{v}(t)\,\mathbf{v}(t-\tau)$$

$$\times f_{0\alpha}(t-\tau)\exp\left[i\omega\tau - ik\int_{t-\tau}^{t} v_x(t')\,dt'\right]. \tag{B8}$$

Near $z=0$, in the inner region $|\partial/\partial z| >> |\partial/\partial x| \sim 1/\lambda$. Then $\underset{\leftrightarrow}{\sigma}$ can be diagonalized in the form

$$\underset{\leftrightarrow}{\sigma} = \begin{pmatrix} \sigma_x & 0 & 0 \\ 0 & \sigma_\tau & 0 \\ 0 & 0 & \sigma_\parallel \end{pmatrix}, \tag{B9}$$

where $\hat{x}=\hat{x}$, $\hat{\tau}=\alpha_z\hat{y}-\alpha_y\hat{z}$, $\hat{b}\equiv\mathbf{B}/B=\alpha_y\hat{y}+\alpha_z\hat{z}$ with $\alpha_y = B_y/B$ and $\alpha_z = B_z/B$. We write $E_{1\tau}=\alpha_z E_{1y}-\alpha_y E_{1z}$, $E_{1\parallel}=\alpha_y E_{1y}+\alpha_z E_{1z}$, since $\alpha_y^2+\alpha_z^2=1$.

Defining $\psi_1\equiv A_{1y}$, $\chi_1\equiv A_{1z}$, and writing, as in WBL, any perturbed quantity such as ψ in the form $\tilde{\psi}(z)\exp(ikx-i\omega t)$, we get

$$-\frac{c}{4\pi}\frac{d^2\tilde{\psi}}{dz^2} \simeq \hat{y}\cdot\underset{\leftrightarrow}{\sigma}\cdot\tilde{\mathbf{E}},$$

$$= \alpha_y\sigma_\parallel\tilde{E}_\parallel + \alpha_z\sigma_\perp\tilde{E}_\tau,$$

$$= (\alpha_y^2\sigma_\parallel + \alpha_z^2\sigma_\perp)\tilde{E}_y + \alpha_y\alpha_z(\sigma_\parallel-\sigma_\perp)\tilde{E}_z, \tag{B10}$$

where $\sigma_\perp \equiv \sigma_x = \sigma_\tau$. Similarly,

$$-\frac{c}{4\pi}\frac{d^2\tilde{\chi}}{dz^2} \simeq \left(\alpha_y^2\sigma_\perp+\alpha_z^2\sigma_\parallel\right)\tilde{E}_z$$

$$+\,\alpha_y\alpha_z(\sigma_\parallel-\sigma_\perp)\tilde{E}_y. \tag{B11}$$

The parallel conductivity $\sigma_\parallel$ can be calculated by using the drift-kinetic approximation for electrons. It can be shown that [*Drake and Lee*, 1977; WBL, 1990]

$$\tilde{J}_\parallel = i\frac{\omega_p^2}{4\pi\omega^2}(\omega-\omega^*)s^2Z'(s)\,\tilde{E}_\parallel, \tag{B12}$$

where

$$\sigma_\parallel = i\frac{\omega_p^2}{4\pi\omega^2}(\omega-\omega^*)s^2Z'(s), \tag{B13}$$

where $\omega^* = kV_{ex}$, $s=\omega/(k_\parallel v_{te})$, and $Z(s)$ is the plasma dispersion function.

The perpendicular conductivity $\sigma_\perp$ is mostly due to the ion polarization drift, since the $\boldsymbol{E}\times\boldsymbol{B}$ drift carries no current. The full ion response [*Cowley et al.* 1986; WBL] essentially reduces to

$$\tilde{\mathbf{J}}_\perp \simeq \frac{nm_ic^2}{B^2}(-i\omega\tilde{\mathbf{E}}_\perp) \simeq -i\omega\frac{c^2}{4\pi V_A^2}\tilde{\mathbf{E}}_\perp, \tag{B14}$$

where $V_A = B/(4\pi n m_i)^{1/2}$ is the Alfvén velocity. Note that

$$\frac{\sigma_\perp}{\sigma_\parallel} \sim \frac{c^2}{V_A^2}\,\frac{\omega^2}{\omega_p^2 s^2}$$

$$= \frac{k_\parallel^2 v_{te}^2}{k_0^2 V_A^2} << \frac{k^2}{k_0^2}\,\frac{m_i}{m_e} \sim 10^{-2} << 1, \tag{B15}$$

for magnetotail plasmas. We then recover the result derived by WBL, i.e.,

$$\frac{d^2}{dz^2}\begin{pmatrix}\tilde{\psi}\\ \tilde{\chi}\end{pmatrix}$$

$$\simeq -\frac{4\pi i\omega}{c^2}\sigma_\parallel\begin{pmatrix}\alpha_y^2 & \alpha_y\alpha_z\\ \alpha_y\alpha_z & \alpha_z^2\end{pmatrix}\begin{pmatrix}\tilde{\psi}\\ \tilde{\chi}\end{pmatrix}, \tag{B16}$$

We now introduce two characteristic frequencies. One is $\omega_e \equiv k_\parallel v_{te}$. The other is $\omega_\varepsilon \equiv \tau_\varepsilon^{-1}$, where τ_ε is the time it takes for an electron to travel a characteristic distance k^{-1} along x [WBL]. From the field-line equation

$$\frac{dz}{B_z} = \frac{dx}{B_x}, \tag{B17}$$

it follows that if an electron travels k^{-1} along x, it must travel z_0 along z, where z_0 is given by the relation

$$z_0^2 = \lambda^2[1-\exp(2\varepsilon/k\lambda)]. \tag{B18}$$

We then obtain [WBL]

$$\tau_\varepsilon = \int_0^{k^{-1}} \frac{dx}{\tilde{v}_{ex}} = \int_0^{z_0} \frac{dx}{\tilde{v}_{ez}},$$

$$\simeq \frac{\pi}{2\varepsilon v_{te}} \exp(\varepsilon/k\lambda). \quad \text{(B19)}$$

We introduce the strained inner variable

$$\xi = \frac{z}{\delta_e} \equiv \frac{1}{s}, \quad \text{(B20)}$$

where $\delta_e \equiv (\omega/k v_{te})(L_s/s_0)$ and $s_0 = 1$ if $\omega_e > \omega_\varepsilon$ (as in region 3), but $s_0 = \omega_e/\omega_e$ if $\omega_e < \omega_\varepsilon$ (as in region 2). The shear length L_s is defined by $L_s \equiv \lambda B(z=0)/B_0$. With these definitions, (B16) can be rewritten as

$$\frac{d^2}{dz^2}\begin{pmatrix}\tilde{\psi}\\ \tilde{\chi}\end{pmatrix} = f(\xi)\begin{pmatrix}\alpha_0^2 & \alpha_0\varepsilon_0\\ \alpha_0\varepsilon_0 & \varepsilon_0^2\end{pmatrix}\begin{pmatrix}\tilde{\psi}\\ \tilde{\chi}\end{pmatrix}, \quad \text{(B21)}$$

where $\alpha_0 \equiv B_y(0)/B_0$, $\varepsilon_0 \equiv B_z(0)/B_0$, and

$$f(\xi) \equiv \frac{(1-\omega^*/\omega)\delta_e^2 \xi^{-2} k_0^2}{\alpha_0^2+\varepsilon_0^2} Z'(\xi^{-1}), \quad \text{(B22)}$$

which is the same as (77) of WBL.

Equation (B21) can be diagonalized as

$$\frac{d^2}{d\xi^2}\begin{pmatrix}a_1\\ a_2\end{pmatrix} = f(\xi)\begin{pmatrix}\alpha_0^2+\varepsilon_0^2 & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}a_1\\ a_2\end{pmatrix}, \quad \text{(B23)}$$

where

$$a_1 = \alpha_0\tilde{\psi} + \varepsilon_0\tilde{\chi}, \quad \text{(B24)}$$

$$a_2 = -\varepsilon_0\tilde{\psi} + \alpha_0\tilde{\chi}. \quad \text{(B25)}$$

As shown by WBL, $a_2(\chi=0) = 0$, which yields

$$\tilde{\psi}(0) = (\alpha_0/\varepsilon_0)\tilde{\chi}(0). \quad \text{(B26)}$$

The solution for a_1 is different in different regions.

In region 1, where $\gamma\tau_b \ll 1$, the integration over z involves a bounce average which extends over the outer region. We can then see the stabilizing effect of this bounce by noting that it tends to cancel out the perturbed current in the inner region and gives $\tilde{\psi} = \tilde{\chi} = 0$. In regions 2 and 3, however, we have $\gamma\tau_b \gtrsim 1$ and $\gamma\tau_b \gg 1$, respectively. Integrating (B23) over the inner region, and using the "constant ψ" approximation, we get the jump condition

$$\frac{da_1}{d\xi}\Big|_{\text{inner}} \simeq (\alpha_0^2+\varepsilon_0^2)\, a_1(0) \int_{\text{inner}} d\xi f(\xi). \quad \text{(B27)}$$

In order to match the inner region to the outer region, we use the relations

$$\Delta_1' \equiv \frac{d}{dz}\log\tilde{\psi}_{\text{outer}}\Big|_{0-}^{0+} = \frac{d}{dz}\log\tilde{\psi}_{\text{outer}}\Big|_{-\infty}^{+\infty}, \quad \text{(B28)}$$

$$\Delta_2' \equiv \frac{d}{dz}\log\tilde{\chi}_{\text{outer}}\Big|_{0-}^{0+} = \frac{d}{dz}\log\tilde{\chi}_{\text{innerr}}\Big|_{-\infty}^{+\infty} \quad \text{(B29)}$$

The left-hand side of (B27) can then be written as

$$\frac{da_1}{d\xi}\Big|_{\text{inner}} = \delta_e \frac{da_1}{d\xi_{\text{inner}}}\Big|_{-\infty}^{+\infty} = \delta_e \frac{da_1}{d\xi_{\text{outer}}}\Big|_{-\infty}^{+\infty}$$

$$= \delta_e\left[\alpha_0\tilde{\psi}(0)\Delta_1' + \varepsilon_0\tilde{\chi}(0)\Delta_2'\right]. \quad \text{(B30)}$$

The right-hand side of (B27) gives

$$\int_{\text{inner}} d\xi f(\xi) = \int_{-\infty}^{+\infty} d\xi f(\xi) = 2\int_{\infty}^{0} ds \frac{d\xi}{ds}\Big|_{\xi=1/s} f(s^{-1})$$

$$= -\frac{2\delta_e^2 k_0^2(1-\omega^*/\omega)}{\alpha_0^2+\varepsilon_0^2} Z(0). \quad \text{(B31)}$$

Using (B30) and (B31), we obtain the dispersion equation

$$D(k,\omega) = \frac{k v_{te}\Delta' S_0}{2k_0^2 L_s \omega} + \left(1-\frac{\omega^*}{\omega}\right)Z(0) = 0, \quad \text{(B32)}$$

where

$$\Delta' \equiv \frac{\alpha_0^2\Delta_1' + \varepsilon_0^2\Delta_2'}{\alpha_0^2+\varepsilon_0^2}. \quad \text{(B33)}$$

Equation (B32) is slightly different from the analogous equation derived by WBL. The difference can be attrributed primarily to a slightly different treatment of the asymptotic matching condition; here we have extended the outer solution all the way to $z=0$ instead of taking it to the outer limits of the inner region. In the rotation of WBL, $\Delta' = \Delta_r' + i\Delta_i'$; however, if we recall that $\Delta_i' \sim \varepsilon \ll \Delta_r'$, it is clear that the results derived by WBL and here do not differ significantly. (In order to avoid possible confusion, we point out that ω^* has been neglected in certain results derived in WBL; this affects the real part of ω, but not the growth rate.) Solving the dispersion equation (B32), we obtain ω_R and γ, given, respectively, by (103) and (104) in region 2 and (105) and (106) in region 3.

The growth rate derived from kinetic theory can also be obtained from the fluid model. The fluid analog of (106) has already been derived in section 3. Here we derive easily the anolog of (104). Recalling the heuristic discussion given by WBL (see section 3, WBL), we have $\gamma \approx \omega_\varepsilon \equiv \omega_\varepsilon k_\parallel v_{te}/\omega_e = \omega_\varepsilon k d_e/(\lambda\omega_e)$. Here $d_e = \Delta_r'/2k_0^2$, which is obtained from (68), with Δ_r' given by (88). If we now use (B25), we get (110) which is larger than the growth rate derived from kinetic theory by the factor $\pi^{1/2}$.

Acknowledgments. We thank A. T. Y. Lui for encouraging us to persevere with this problem, and for asking searching questions. We are also grateful to D. G. Sibeck for a useful discussion. This research is supported by the National Science Foundation grants ATM–9100513, ATM–9209006 and the Air Force Office of Scientific Research grant F49620-93-1-0071.

The Editor thanks J. Buechner and another referee for their assistance in evaluating this paper.

REFERENCES

Akasofu, S. -I., A. T. Y. Lui, C. I. Meng, and M. Haurwitz, Need for a three-dimensional analysis of magnetic fields in the magnetotail during substorms, *Geophys. Res. Lett., 5*, 283, 1978.

Ambrosiano, J., and L. C. Lee, Simulation of the ion tearing mode in the presence of a background plasma, *J. Geophys. Res., 88*, 7860, 1983.

Ambrosiano, J., L. C. Lee, and Z. F. Fu, Simulation of the collisionless tearing instability in an anisotropic neutral sheet, *J. Geophys. Res., 91*, 113, 1986.

Bieber, J. W., E. C. Stone, E. W. Hones, Jr., D. N. Baker, S. J. Bame, and R. P. Lepping, Microstructure of magnetic reconnection in Earth's magnetotail, *J. Geophys. Res., 89*, 6705, 1984.

Birn, J., R. R. Sommer, and K. Schindler, Open and closed magnetospheric tail configurations and their stability, *Astrophys. Space Sci., 35*, 389, 1975.

Birn, J., Self-consistent magnetotail theory: general solution for the quiet tail with vanishing field aligned currents, *J. Geophys. Res., 84*, 5143, 1979.

Büchner, J., and L. M. Zelenyi, Chaotization of the electron motion as the cause of an internal magnetotail instability and substorm onset, *J. Geophys. Res., 92*, 13456, 1987.

Büchner, J., M. Kuznetsova, and L. M. Zelenyi, Sheared field tearing mode instability and creation of flux ropes in the Earth magnetotail, *Geophys. Res. Lett., 18*, 385, 1991.

Cattell, C. A., and F. S. Mozer, Electric fields measured by ISEE 1 within and near the neutral sheet during quiet and active times, *Geophys. Res. Lett., 9*, 1041, 1982.

Chen, J., and P. J. Palmadesso, Chaos and nonlinear dynamics of single-particle orbits in a magnetotail-like magnetic field, J. Geophys. Res. 91, 1499 (1986).

Coppi, B., G. Laval, and R. Pellat, Dynamics of the geomagnetic tail, *Phys. Rev. Lett., 19*, 1207, 1966.

Coroniti, F. V., On the tearing mode in quasi-neutral sheets, *J. Geophys. Res., 85*, 6719, 1980.

Cowley, S. C., R. M. Kulsrud, and T. S. Hahm, Linear stability of tearing modes, *Phys. Fluids, 29*, 3230, 1986.

Drake, J. F., and Y. C. Lee, Kinetic theory of tearing instabilities, *Phys. Fluids, 20*, 1341, 1977.

Fairfield, D. H., On the average configuration of the geomagnetic tail, *J. Geophys. Res., 84*, 1950, 1979.

Furth, H. P., The 'mirror instability' for finite particle gyro-radius, *Nucl. Fusion Suppl.*, part 1, 169, 1962.

Furth, H. P., Hydromagnetic instabilities due to finite resistivity, in *Propagation and Instabilities in Plasmas*, edited by W. I. Futterman, Stanford University Press, Stanford, Calif., 1963.

Furth, H. P., Instabilities due to finite resistivity or finite current-carrier mass, in *Advanced Plasma Theory*, Proceedings of the International School of Physics "Enrico Fermi," Course XXV, edited by M. N. Rosenbluth, p. 159, Academic, San Diego, Calif., 1964.

Furth, H. P., J. Killeen, and M. N. Rosenbluth, Finite-resistivity instabilities of a sheet pinch, *Phys. Fluids, 6*, 459, 1963.

Galeev, A. A., in *Basic Plasma Physics*, vol. 2, edited by A. A. Galeev and R. N. Sudan, p. 305, North-Holland, New York, 1984.

Galeev, A. A., and L. M. Zelenyi, Tearing instability in plasma configurations, *Sov. Phys. JETP, 43*, 1113, 1976.

Greene, J. M., Geometrical properties of three-dimensional reconnecting magnetic fields with nulls, *J. Geophys. Res., 93*, 8583, 1988.

Hamilton, J. E. M., and J. W. Eastwood, The effect of a normal magnetic field component on current sheet stability, *Plasma Phys., 39*, 293, 1982.

Hau, L. N., and G. -H. Voigt, Loss of MHD equilibrium caused by the enhancement of the magnetic B_y component in Earth's magnetotail, *J. Geophys. Res., 97*, 8707, 1992.

Horton, W., and T. Tajima, Decay of correlations and the collisionless conductivity in the geomagnetictail, *Geophys. Res. Lett., 17*, 123, 1990.

Krall, N. A., and A. W. Trivelpiece, *Principles of Plasma Physics*, p. 95, San Francisco Press, San Francisco, Calif., 1986.

Kruskal, M. D., and C. R. Oberman, On the stability of plasma in static equilibrium, *Phys. Fluids, 1*, 275, 1958.

Kuznetsova, M. M., and L. M. Zelenyi, Magnetic reconnection in collisionless field reversals: The universality of the ion tearing mode, *Geophys. Res. Lett., 18*, 1825, 1991.

Lau, Y. -T., and J. M. Finn, Three-dimensional kinematic reconnection in the presence of field nulls and closed field lines, *Ap. J., 350*, 672, 1990.

Laval, G., R. Pellat, and M. Vuillemin, Instabilités electromagnétique des plasmas sans collisions, in *Plasma Physics and Controlled Fusion Research*, Vol. II, p. 259, International Atomic Energy Agency, Vienna, 1966.

Lembege, B., and R. Pellat, Stability of a thick two-dimensional quasineutral sheet, *Phys. Fluids, 25*, 1995, 1982.

Lepping, R. P., Dialog on the phenomenological model of substorms in the magntotail, in *Magnetotail Physics*, edited by A. T. Y. Lui, p. 419, Johns Hopkins University Press, Baltimore, Md., 1987.

Lopez, R. E., A. T. Y. Lui, D. G. Sibeck, K. Takahashi, R. W. McEntire, L. J. Zanetti, and S. M. Krimigis, On the relationship between the energetic particle flux morphology and the change in the magnetic field magnitude during substorms, *J. Geophys. Res., 94*, 17,105, 1989.

Lui, A. T. Y., Characteristics of the cross-tail current in the Earth's magnetotail, in *Magnetospheric Currents, Geophys. Monogr. Ser.*, vol. 28, edited by T. Potemra, p. 158, AGU, Washington, D. C., 1983.

Lui, A. T. Y., Roadmap to magnetotail domains, in *Magnetotail Physics*, edited by A. T. Y. Lui, p. 3, Johns Hopkins University Press, Baltimore, Md., 1987.

Lui, A. T. Y., R. E. Lopez, S. M. Krimigis, R. W. McEntire, L. J. Zanetti, and T. A. Potemra, A case study of magnetotail current disruption and diversion, *Geophys. Res. Lett.,15*, 721, 1988.

McComas, D. J., C. T. Russell, R. C. Elphic, and S. J. Bame, The near-earth cross-tail current sheet: Detailed ISEE-1 and ISEE-2 case studies, *J. Geophys. Res., 91*, 4287, 1986.

Ness, N. F., The Earth's magnetic tail, *J. Geophys. Res., 70*, 2989, 1965.

Nishida, A., Y. K. Tulunay, F. S. Mozer, C. A. Cattell, E. W. Hones, Jr., and J. Birn, Electric field evidence for tailward flow at substorm onset, *J. Geophys. Res., 88*, 9109, 1983.

Paranicas, C., and A. Bhattacharjee, Relaxation of magnetotail plasmas with field-aligned currents, *J. Geophys. Res., 94*, 479, 1989.

Pellat, R., F. V. Coroniti, and P. L. Pritchett, Does ion tearing exist?, *Geophys. Res. Lett., 18*, 143, 1991.

Pritchett, P. L., F. V. Coroniti, R. Pellat, and H. Karimabadi, Collisionless reconnection in a quasi-neutral sheet near marginal stability, *Geophys. Res. Lett., 16*, 1269, 1989.

Pritchett, P. L., F. V. Coroniti, R. Pellat, and H. Karimabadi, Collisionless reconnection in two-dimensional magnetotail equilibria, *J. Geophys. Res., 96*, 11,523, 1991.

Schindler. K., A theory of the substorm mechanism, *J. Geophys. Res., 79*, 2803, 1974.

Sibeck, D. G., G. L. Siscoe, J. A. Slavin, E. J. Smith, B. T. Tsurutani, and R. P. Lepping, The distant magnetotail's response to a strong interplanetary magnetic field B_y: Twisting, flattening and field-line bending, *J. Geophys. Res., 90*, 4011, 1985.

Swift, D. W., A two-dimensional simulation of the interaction of the plasma sheet with the lobes of the Earth's magnetotail, *J. Geophys. Res., 88*, 125, 1983.

Swift, D. W., and C. Allen, Interaction of the plasma sheet with the lobes of the Earth's magnetotail, *J. Geophys. Res., 92*, 10015, 1987.

Takahashi, K., L. J. Zanetti, R. E. Lopez, R. W. McEntire, T. A. Potemra, and K. Yumoto, Disruption of the magnetotail current sheet observed by AMPTE/CCE, *Geophys. Res. Lett., 14*, 1019, 1987.

Terasawa, T., Numerical study of explosive tearing mode instability in one-component plasmas, *J. Geophys. Res., 86*, 9007, 1981.

Tsurutani, B. T., J. A. Slavin, E. J. Smith, R. Okida, and D. E. Jones, Magnetic structure of the distant geotail from 60 to 220 R_E: ISEE-3, *Geophys. Res. Lett., 11*, 1, 1984.

Voigt, G. -H., and R. V. Hilmer, The influence of the IMF B_y component on the Earth's magneto-hydrostatic magnetotail, in *Magnetotail Physics*, edited by A. T. Y. Lui, p. 91, Johns Hopkins University Press, Baltimore, Md., 1987.

Wang, X., A. Bhattacharjee, and A. T. Y. Lui, Collisionless tearing instability in magnetotail plasmas, *J. Geophys. Res., 95*, 15047, 1990.

Wang, X., A. Bhattacharjee, and A. T. Y. Lui, Theory of current disruption and diversion in magnetotail plasmas: Recent developments, in the *Physics of Space Plasmas (1990)*, S.P I Conf. Proc. Reprint Ser., vol. 10, edited by T. Chang, G. B. Crow and J. R. Jasperse, p. 162, Scientific, Cambridge, Massa., 1991.

Zwingmann, W., Self-consistant magnetotail theory: Equilibrium structures including arbitrary variation along the tail axis, *J. Geophys. Res., 88*, 9101, 1983.

Zwingmann, W., and K. Schindler, Magnetic islands in the quiet magnetotail, *Geophys. Res. Lett., 7*, 909, 1980.

Zwingmann, W., J. Wallace, K. Schindler, and J. Birn, Particle simulation of magnetic reconnection in the magnetotail configuration, *J. Geophys. Res., 95*, 20,877, 1990.

A. Bhattacharjee and X. Wang, Department of Applied Physics, Columbia University, 202 S. W. Mudd Bldg., 500 W. 120th Street, New York, NY 10027.

(Received October 5, 1992;
revised April 21, 1993;
accepted March 23, 1993.)

Tearing stability of the two-dimensional magnetotail

B. G. Harrold,[a] A. Bhattacharjee, and Xiaogang Wang
Department of Physics and Astronomy, University of Iowa, Iowa City, Iowa 52242

(Received 31 March 1995; accepted 11 July 1995)

The linearized incompressible magnetohydrodynamic equations that include a generalized Ohm's law are solved for tearing eigenmodes of a plasma sheet with a normal magnetic field (B_n). In contrast to the Harris sheet with the equilibrium magnetic field $[\mathbf{B}=B_0 \tanh(z/a)\hat{\mathbf{x}}]$, the two-dimensional plasma sheet with the field $[\mathbf{B}=B_0 \tanh(z/a)\hat{\mathbf{x}}+B_n\hat{\mathbf{z}}]$, in which the B_n field lies in the plane of the B_x field, has no neutral line if $B_n\neq 0$. Such a geometry is intrinsically resilient to tearing because it cannot change topology by means of linear perturbations. This qualitative geometrical idea is supported by calculations of growth rates using a generalized Ohm's law that includes collisional resistivity and finite electron inertia as the mechanisms for breaking field lines. The presence of B_n reduces the resistive tearing mode growth rate by several orders of magnitude (assuming $B_n/B_0\sim 0.1$) compared with that in the Harris sheet model ($B_n=0$). The growth rate scaling with Lundquist number (S) has the typical $S^{-3/5}$ ($S^{-1/3}$) dependence for large (small) wave numbers and very small values of B_n. For larger values of B_n, all modes behave diffusively, scaling as S^{-1}. The collisionless electron tearing mode growth rate is found to be proportional to δ_e^2 in the presence of significant $B_n(>10^{-2}B_0)$ and large $k_x(\sim 0.1a^{-1}-0.5a^{-1})$, and becomes completely stable ($\gamma<0$) for $B_n/B_0\geq 0.2$. Implications for magnetospheric substorms are discussed. © *1995 American Institute of Physics.*

I. INTRODUCTION

The tearing stability of the Earth's magnetotail has been a subject of considerable interest over the last three decades. The subject is considered to be relevant to the issue of substorm onset that is widely believed to occur at near-Earth distances, at about $10R_E$ or less (where R_E is the radius of the Earth).

It is well known that the resistivity of the magnetotail is so low that the plasma can be regarded as virtually collisionless. In such a plasma, finite particle inertia is a possible mechanism for magnetic reconnection. Hence, most theoretical analyses of the tearing stability of the tail rely on kinetic theory.[1–10] Electromagnetic particle-in-cell (PIC) simulations that follow individual particle dynamics in self-consistent electromagnetic fields have also been developed to study the stability of some simple models of the magnetotail.[11–17] While PIC codes do include, in principle, the relevant physics at the microscopic or kinetic level, these codes cannot be used at the present time to study reliably global three-dimensional dynamics of the magnetosphere. On the other hand, global codes based on resistive magnetohydrodynamics (MHD) have been developed,[18–22] and are routinely used to study the dynamics of the magnetosphere including the magnetotail. These codes generally assume the existence of a large plasma resistivity many orders of magnitude larger than the classical value. Although it is not clear whether the assumption of a large resistivity is justified, global resistive MHD codes have shed valuable light on the overall geometric features of magnetospheric dynamics.

In this paper, we study the linear resistive and collisionless tearing stability of the two-dimensional (2-D) magnetotail. For Ohm's law, we use

$$\mathbf{E}+\frac{\boldsymbol{\nu}\times\mathbf{B}}{c}=\eta\mathbf{J}+\frac{m_e}{ne^2}\frac{d\mathbf{J}}{dt}. \tag{1}$$

Here $\mathbf{E}$ is the electric field, $\mathbf{B}$ is the magnetic field, $\boldsymbol{\nu}$ is the plasma velocity, $\mathbf{J}$ is the current density, η is the plasma resistivity, m_e is the electron mass, e is the electron charge, n is the plasma density, c is the speed of light, and $d/dt=\partial/\partial t+\boldsymbol{\nu}\cdot\nabla$. The Ohm's law (1), which is by no means complete, allows us to examine the tearing stability of the 2-D magnetotail for two mechanisms that can break field lines: resistivity and electron inertia. Despite the simplicity of our Ohm's law, we will demonstrate that it provides insight regarding the conditions under which collisional as well as collisionless tearing modes are possible in the magnetotail.

One of the issues brought out by our study is the important role of magnetic *geometry* on the tearing stability of the tail. We consider the simple model

$$\mathbf{B}=B_0 \tanh(z/a)\hat{\mathbf{x}}+B_n\hat{\mathbf{z}}, \tag{2}$$

where B_n is a constant magnetic field in the plane of the B_x field. If $B_n=0$, the model (2) reduces to the Harris sheet [Fig. 1(a)], which is well known to be unstable to resistive and collisionless tearing modes. We have known this for so long and with such certainty that we rarely find it necessary to emphasize that the reason why the Harris sheet is readily amenable to tearing is because when $B_n=0$, we already have a neutral line at $z=0$. In the presence of this neutral line, a linear, periodic magnetic perturbation of the form, say, $b\sin(kx)\hat{\mathbf{z}}$, produces neighboring equilibria with magnetic islands with lower energy than the initial state [Fig. 1(b)]. The geometry of the Harris sheet with a clearly defined neutral line favors topological change, and the tearing instability provides a mechanism by which the plasma can have access

[a]Present address: Department of Physics, University of Alberta, Edmonton, Alberta, T6G 2J1, Canada.

 1070-664X/95/2(10)/3857/8/$6.00 © 1995 American Institute of Physics

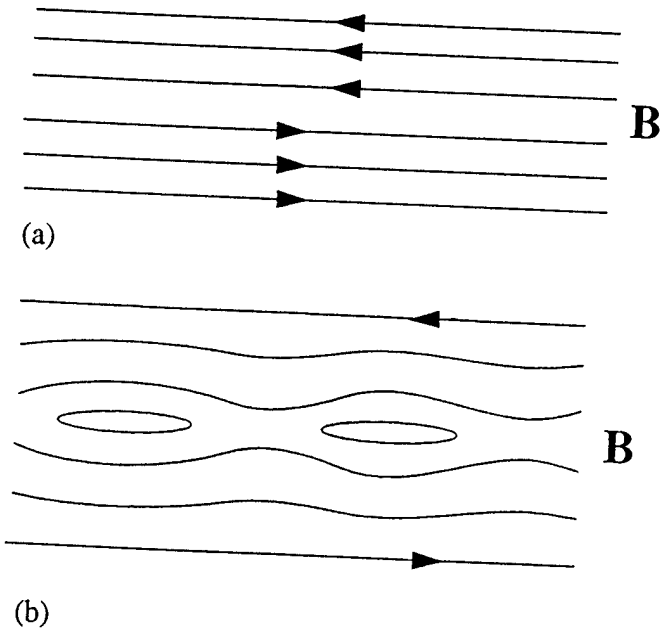

FIG. 1. (a) The topology of field lines in the Harris equilibrium $\mathbf{B}=B_0 \tanh(z/a)\hat{\mathbf{x}}$. (b) The topology of field lines when the perturbation $\mathbf{b}=b \sin(kx)\hat{\mathbf{z}}$ is imposed on the Harris equilibrium.

to a lower-energy state. In the presence of the resistive term on the right-hand side of (1), Furth, Killeen, and Rosenbluth (hereafter referred to as FKR)[23] showed that there is a tearing instability that has a growth rate proportional to $\eta^{3/5}$. For small values of η, it is much faster than the time scale of resistive diffusion (which is proportional to η).

What we have discussed in the last paragraph is obvious and well known, but it bears repetition because it allows us to contrast the geometric differences between the cases $B_n=0$ and $B_n\neq 0$. If $B_n\neq 0$, the field lines become parabolic, and the neutral line in the Harris sheet is eliminated (Fig. 2). A linear perturbation of the form $b \sin(kx)\hat{\mathbf{z}}$ does not affect genuine topological change and create islands. Instead, what we have is a neighboring equilibrium with a topology similar to the initial equilibrium (indicated by dashed and solid lines, respectively, in Fig. 2). The main difference between the two neighboring equilibria in Fig. 2 is the quasiperiodic bunching and unbunching of field lines. We claim that the resilience of the geometry depicted in Fig. 2 to topological change is enough to raise serious doubts about the possibility of exciting robust linear tearing instabilities in the magnetotail (unless B_n is very small or zero). By robust instabilities, we mean those that can be distinguished sufficiently from mere diffusion. It is extremely unlikely that feeble tearing instabilities that are virtually indistinguishable from diffusion and do not induce genuine topological change can account for the rapid temporal signatures associated with substorms.

The ideas discussed above are borne out qualitatively by the linear stability calculations presented in this paper. In the discussion above, we have chosen to emphasize geometry over dynamics, because we feel that simple considerations of geometry are helpful in developing intuition about the conditions under which we can expect tearing instabilities to be interesting candidates for substorms.

There has been a tendency in the magnetotail literature on this problem to emphasize the differences in particle orbits (such as intrinsic chaoticity) between the equilibria with and without B_n. While we do not deny these important differences, we believe that an overemphasis on details of particle orbits have detracted attention from the simpler and more important geometrical issue stressed in this paper. Our results suggest that B_n has a strongly stabilizing effect on all tearing modes, independent of the mechanism that breaks field lines. In this sense, our conclusions are more general than those of Pellat *et al.*[8] who have questioned the existence of the ion tearing instability in the parabolic equilibrium (2).

The following is a plan of our paper. In Sec. II, we give the linearized eigenmode equations for resistive and collisionless tearing stability of a model magnetotail equilibrium. In Secs. III and IV, we present numerical results on resistive and collisionless stability, respectively, establishing connections with known analytical results wherever possible. In Sec. V, we conclude with a summary and the implications of our results for substorm phenomena.

II. METHOD OF ANALYSIS

We use a slab model for the ambient plasma and magnetic field, and analyze linear stability of the resistive and electron tearing modes in the magnetic field (2). In equilibrium, the normal magnetic field B_n gives rise to a force $\mathbf{f}=-\rho_0(z)g\hat{\mathbf{x}}$, where g is an effective gravitational acceleration and $\rho_0(z)=\rho_0(0)\mathrm{sech}^2(z/a)$. The relevant linear single-fluid MHD equations are

$$\frac{\partial \rho}{\partial t}+\nabla\cdot(\rho_0 \boldsymbol{\nu})=0, \tag{3}$$

$$\rho_0 \frac{\partial \boldsymbol{\nu}}{\partial t}=-\nabla p+\frac{1}{4\pi}(\nabla\times\mathbf{b}\times\mathbf{B}+\nabla\times\mathbf{B}\times\mathbf{b})-\rho g\hat{\mathbf{x}}, \tag{4}$$

$$\frac{\partial \mathbf{b}}{\partial t}=\nabla\times(\boldsymbol{\nu}\times\mathbf{B})-\nabla\times\left(\frac{c^2}{\omega_{pe}^2}\frac{\partial}{\partial t}\nabla\times\mathbf{b}+\frac{c^2\eta}{4\pi}\nabla\times\mathbf{b}\right), \tag{5}$$

where equilibrium variables are designated by the subscript zero, and linearized dependent variables are without sub-

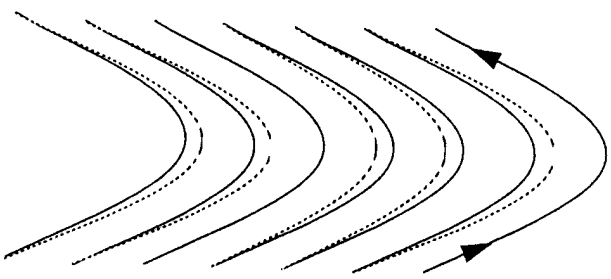

FIG. 2. The topology of field lines when the perturbation $\mathbf{b}=b \sin(kx)\hat{\mathbf{z}}$ is imposed on the equilibrium $\mathbf{B}=B_0 \tanh(z/a)\hat{\mathbf{x}}+B_n\hat{\mathbf{z}}$.

script. For simplicity, we consider incompressible perturbations. The analysis of Wang and Bhattacharjee[24] clearly demonstrates that fluid compressibility has a strongly stabilizing effect, and the assumption of incompressibility, therefore, biases this analysis in favor of instability. By the same token, if one finds stability by this analysis, there is no doubt that the full dynamics including compressibility effects will yield an even larger parameter regime over which the plasma sheet is stable. We represent the two-dimensional (incompressible) perturbed velocity and magnetic fields by $\boldsymbol{\nu}=\nabla\times(\phi\hat{\mathbf{y}})$ and $\mathbf{b}=\nabla\times(A\hat{\mathbf{y}})$, respectively, with $\partial/\partial y=0$ and $\partial/\partial x=ik_x$. Assuming linearized perturbations of the form $\exp(ik_x x-i\omega t)$, we obtain the following system of equations of for ϕ and A:

$$r\frac{\partial^2\phi}{\partial z^2}=p\frac{\partial\phi}{\partial z}+q\phi+uA+\frac{\partial A}{\partial z}, \tag{6}$$

$$\frac{\partial^2 A}{\partial z^2}=F\frac{\partial\phi}{\partial z}+G\phi+EA. \tag{7}$$

Here we have used the following definitions:

$$F=-b_n\rho\left(\frac{\rho}{S}-i\omega\delta_e^2\right)^{-1},\quad G=-ik_xB\rho\left(\frac{\rho}{S}-i\omega\delta_e^2\right)^{-1},$$

$$E=k_x^2-i\omega\rho\left(\frac{\rho}{S}-i\omega\delta_e^2\right)^{-1},$$

$$S\equiv\frac{\tau_R}{\tau_A},\quad \delta_e^2\equiv\frac{c^2}{\omega_p^2(0)a^2},\quad \tau_R\equiv\frac{4\pi a^2}{\eta c^2},$$

$$\tau_A\equiv a\frac{\sqrt{4\pi\rho_0(0)}}{B_\infty},$$

$$r=\left(\frac{\rho}{S}-i\omega\delta_e^2\right)\frac{1}{b_n}+i\frac{b_n}{\omega},$$

$$u=\frac{k_x}{\omega b_n}\left\{\left(\frac{1}{S}-\frac{i\omega\delta_e^2}{\rho}\right)\frac{\partial^2 B}{\partial z^2}+i\omega B+\frac{b_n\omega}{k_x}\times\left(\frac{1}{S}-\frac{i\omega\delta_e^2}{\rho}\right)\frac{\partial}{\partial z}\left[\left(\frac{1}{S}-\frac{i\omega\delta_e^2}{\rho}\right)^{-1}\right]\right\},$$

$$q=\left\{\frac{k_x^2}{b_n}+\frac{ik_x}{\omega^2\rho}\frac{\partial^2\rho}{\partial z^2}+\frac{ik_x^2B^2}{b_n\omega(\rho/S-i\omega\delta_e^2)}+\frac{k_x}{\rho\omega}\frac{\partial}{\partial z}\left[B\rho\left(\frac{\rho}{S}-i\omega\delta_e^2\right)^{-1}\right]\right\}\left(\frac{\rho}{S}-i\omega\delta_e^2\right),$$

$$p=\left[\frac{ib_nk_x}{\omega^2}-1-b_n^2\delta_e^2\left(\frac{\rho}{S}-i\omega\delta_e^2\right)^{-2}\right]\left(\frac{\rho}{S}-i\omega\delta_e^2\right)\times\frac{1}{b_n\rho}\frac{\partial\rho}{\partial z}+2\frac{k_xB}{\omega},$$

and the following normalizations:

$$\frac{y}{a}\to y,\quad \frac{\rho_0}{\rho_0(0)}\to\rho(z),\quad \frac{A}{aB_\infty}\to A(z),$$

$$\frac{t}{\tau_A}\to t,\quad k_xa\to k_x,\quad \frac{\phi\tau_A}{a^2}\to\phi(z),$$

$$\omega\tau_A\to\omega,\quad \frac{B_{x0}}{B_\infty}\to B(z),\quad \frac{B_n}{B_\infty}\to b_n.$$

Equilibrium pressure balance dictates that $g=ab_n/\tau_A^2$.

Equations (6) and (7) generalize the equations given in Ref. 25 by including the effect of the electron inertia term in Ohm's law. We investigate normal modes with tearing parity. At $z=0$, we use the boundary conditions $\phi=0$ and $dA/dz=0$. As $z\to\infty$, the coefficients in (6) and (7) become constants and we incorporate evanescent conditions $\phi(z\to\infty)\sim C\exp(-k_xz)+D\exp(-\kappa z)$ [where κ is a complex constant with a positive real part that is obtained from (6) and (7)], and $A(z\to\infty)\sim\exp(-k_xz)$.

Equations (6) and (7) are converted to a single fourth-order differential equation in ϕ, which is then solved by method of collocation with seventh-order splines.[25] Comparison with previously published numerical results[26] for the resistive tearing mode indicates eigenvalues ω of six digit accuracy.

III. RESISTIVE TEARING

The resistive tearing mode is examined with $\delta_e=0$ in the eigenmode equations. The analytical results of FKR and Coppi *et al.*[27] for the tearing mode growth rate in the Harris sheet (for the "constant-ψ" and "non-constant-ψ" modes, respectively) are as follows:

$$\gamma\sim\begin{cases}S^{-3/5}k_x^{-2/5}(1-k_x^2)^{2/5}, & k_xS^{1/4}\gg1,\\ S^{-1/3}k_x^{2/3}, & k_xS^{1/4}\ll1,\quad k_x\ll1.\end{cases} \tag{8}$$

The growth rates obtained from the numerical solutions of Eqs. (6) and (7) ($b_n=B_n/B_0=0$) are compared to the analytic expressions (8) in Fig. 3. The numerical results shown in Fig. 3(a) clearly exhibit the transition from "non-constant-ψ" modes to "constant-ψ" modes at $k_xS^{1/4}\sim1$, consistent with the predictions of analytical theory. In Fig. 3(b), we compare the analytical and numerical growth rates for a "constant-ψ" FKR mode. The mode growth rate continues to scale as $S^{-3/5}$ accurately in the high-S ($>10^8$) regime.

We now examine the consequence of adding a B_n field to the Harris equilibrium, i.e., the equilibrium magnetic field is now given by (2). For application to the Earth's plasma sheet (where a is or the order $1R_E$ and the system size along x is of the order 10^2R_E), physically relevant values of normalized k_x obey the inequality $k_x\gtrsim0.1$. Since marginal stability without B_n occurs at (normalized) $k_x=1$, we will examine modes with $k_x=0.5$, unless otherwise indicated. Figure 4 shows the spatial profiles of A and ϕ for values of the Lundquist number $S=10^8$ and the normal field component $b_n=10^{-2}$. The growth rate of this mode is $\gamma\tau_A=1.28\times10^{-7}$, which is an extremely weak instability.

In Fig. 5, we show the growth rate as a function of b_n for representative values of S. We see that the small growth rate of the mode in Fig. 4 is typical for a system with $b_n>10^{-3}$, and that the growth rate decreases rapidly with increasing b_n. Inspection of the curve for $S=10^{10}$ shows that, for intermediate values of b_n, the growth rate scales as $b_n^{-2/3}$, while for larger values it scales as b_n^{-1}.

Figure 6 shows the dependence of the mode growth rate on the Lundquist number S for various values of b_n. For this

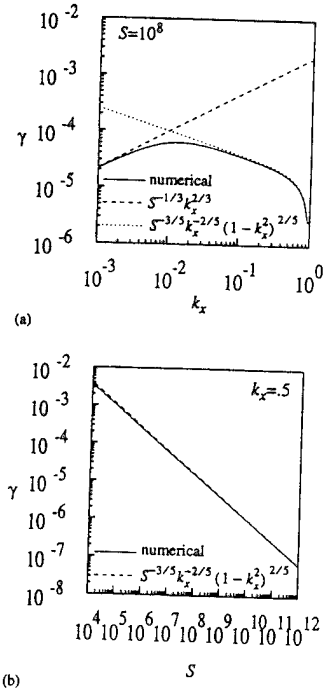

FIG. 3. (a) Comparison of analytic and numerical estimates of the tearing mode growth rate as a function of wave number for the Harris equilibrium. (b) Comparison of analytic and numerical estimates of the tearing mode growth rate as a function of Lundquist number for the Harris equilibrium.

wave number ($k_x=0.5$) and $b_n=10^{-5}$, there is a transition from an $S^{-3/5}$ to an S^{-1} dependence that occurs near $S\sim10^8$. The transition occurs at progressively smaller Lundquist numbers as b_n increases. In contrast, for the Harris equilibrium ($b_n=0$), the mode continues to persist with the $S^{-3/5}$ scaling for all values of S in the high-Lundquist-number regime ($S>10^6$) [see Fig. 3(b)]. The transition to the S^{-1} regime (in which the growth rate is proportional to η, not a fractional power of η) for $b_n\neq0$ was reported earlier by Nishikawa and Sakai,[25] and can be understood analytically from the calculations given in Refs. 28 and 29. To see this, we note that for constant-temperature plasma, Eq. (3) controls the dynamics for the perturbed pressure (if we multiply the perturbed density by a constant temperature). Due to the presence of b_n in the equilibrium, finite pressure gradients, proportional to b_n, must be included in the FKR analysis. The equation for the perturbed pressure is

$$\frac{\partial p_1}{\partial t}+\boldsymbol{v}\cdot\nabla p_0=0, \tag{9}$$

for incompressible displacements, which corresponds to setting $\gamma=0$ in the equations for perturbed pressure in Refs. 28 and 29. After some algebra, it can be shown that the dispersion equation of FKR is modified, and that finite pressure, introduced by the presence of b_n, yields a dispersion equation of the form given by Eq. (15) of Ref. 29. This dispersion equation, in turn, predicts an instability with a growth rate scaling linearly with resistivity (see Sec. III of Ref. 29). If the effect of plasma compressibility is included ($\gamma\neq0$), then, as shown in Ref. 29, the mode growth rate falls to zero faster than S^{-1}.

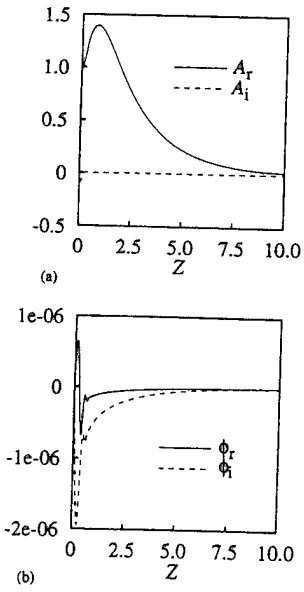

FIG. 4. (a) The eigenmode structure of A (real and imaginary parts) and (b) that of ϕ (real and imaginary parts) for the parameters $k_x=0.5$, $S=10^8$, and $b_n=10^{-2}$.

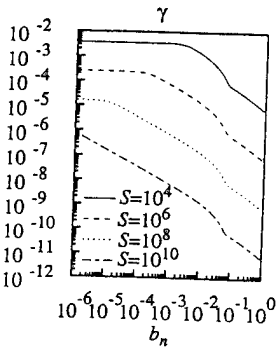

FIG. 5. Dependence of the growth rate on the normal component of the ambient magnetic field and on the Lundquist number.

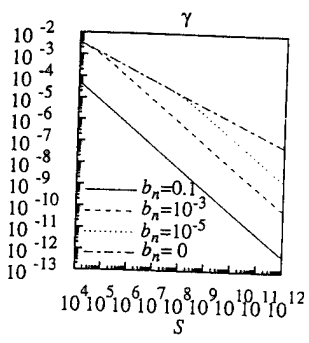

FIG. 6. Dependence of the growth rate on S and b_n for $k_x=0.5$. For $S>10^8$ and b_n nonzero, $\gamma\sim S^{-1}$. For $S<10^8$, $\gamma\sim S^{3/5}$ for small b_n and $\gamma\sim S^{-1}$ for large b_n.

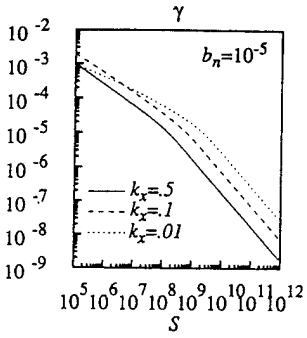

FIG. 7. Dependence of the growth rate on the Lundquist number and wave number for $b_n=10^{-5}$. For $S>10^8$, $\gamma\sim S^{-1}$. For $S<10^8$, $\gamma\sim S^{-3/5}$ for large k_x and $\gamma\sim S^{-1/3}$ for small k_x.

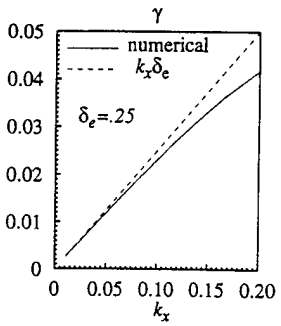

FIG. 8. Comparison of analytic (from Ref. 30) and numerical growth rates [from Eqs. (6) and (7)] for the collisionless tearing mode for the equilibrium magnetic field (10).

We believe that the conversion of an instability that grows as $S^{-3/5}$ to one that grows as S^{-1} and becomes eventually indistinguishable from diffusion, is symptomatic of a deeper geometrical cause. As we have shown, this transition is absent when $b_n=0$, and appears even for an extremely small value of $b_n(=10^{-5})$. We contend that the transition to the η regime (growth rate proportional to S^{-1}) for even a small but positive value of b_n indicates the resilience of equilibrium (2) to the resistive instabilities that are favored when an equilibrium can undergo topological change.

In Fig. 7 we have plotted the growth rate for a few different wave numbers with a very small $b_n(=10^{-5})$. For $10^5<S<10^8$, the growth rate behavior is Harris-like: proportional to $S^{-3/5}$ for $k_x=0.5$, 0.1, and proportional to $S^{-1/3}$ for $k_x=0.01$. However, both long- and short-wavelength modes are in the η regime for sufficiently large Lundquist numbers—a result that has no analog in the Harris equilibrium.

The overall low growth rates for typical plasma sheet parameters, motivates us to look at collisionless effects in Ohm's law. We now report the results when electron inertia [the second term in Eq. (1)] provides the mechanism for breaking field lines.

IV. COLLISIONLESS TEARING

The collisionless tearing mode is examined with $S=\infty$ in the eigenmode equations (6) and (7). We have tested our numerical algorithm with the analytical results of Ottaviani and Porcelli[30] using their equilibrium:

$$\mathbf{B}=B_0\sin(\pi z/a)\hat{\mathbf{x}}+B_y\hat{\mathbf{y}}, \tag{10}$$

with periodic boundary conditions at $z=\pm a$. Equilibrium (10) has magnetic shear, but no normal magnetic field. Ottaviani and Porcelli obtain

$$\gamma\sim k_x\delta_e, \quad 8\delta_e/(k_x^2\pi)\gg 1, \tag{11}$$

where lengths and times are normalized to a/π and $a(4\pi\rho_0)^{1/2}/(B_0\pi)$, respectively. Figure 8 is a comparison of Eq. (11) with our numerical solution of Eqs. (6) and (7) using $\delta_e=0.25$. With this check of the collisionless results in the absence of a normal equilibrium magnetic field completed, we return to the modified Harris sheet equilibrium [Eq. (2)] to examine the effect of a normal magnetic field.

Figure 9 shows the typical spatial profiles of A and ϕ for the value of $\delta_e=10^{-1}$. Comparison with Fig. 4 indicates that there is little difference in the qualitative spatial structure of the perturbed flux. The perturbed streamfunction, however, exhibits fine-scale oscillatory behavior, whereas the resistive ϕ exhibits predominantly exponential decay. In the Appendix, we demonstrate that the wave number of the oscillations along z scales approximately as $k_z\sim k_x/b_n$. This dependence on b_n prevents the accurate determination of numerical solutions for $b_n<10^{-6}$ using our numerical method.

In Fig. 10 we show the growth rate as a function of normal magnetic field b_n for values of the electron inertial length $\delta_e=0.1,0.026$ and wave number $k_x=0.1$, 0.5. There is a nonmonotonic dependence of growth rate on b_n near

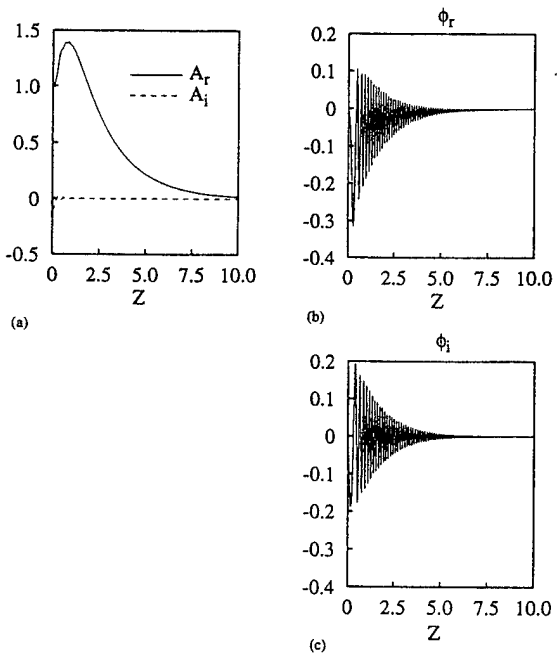

FIG. 9. Eigenmode structure for collisionless tearing with $\delta_e=10^{-1}$, $k_x=0.5$, and $b_n=10^{-2}$. (a) Real and imaginary parts of A. (b) Real part of ϕ. (c) Imaginary part of ϕ.

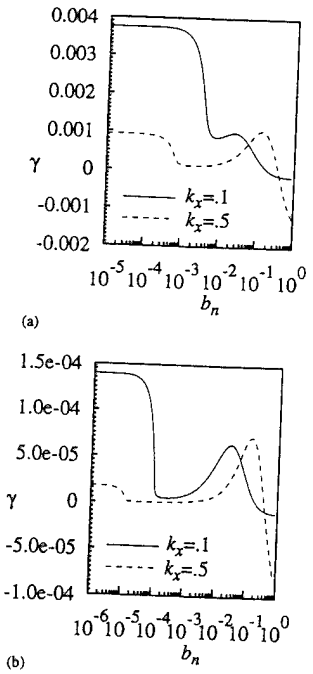

FIG. 10. Collisionless tearing growth rate as a function of b_n for the indicated wave numbers. (a) δ_e=0.1. (b) δ_e=0.026. The plasma sheet in the Earth's magnetotail has $\delta_e \sim 10^{-3}$.

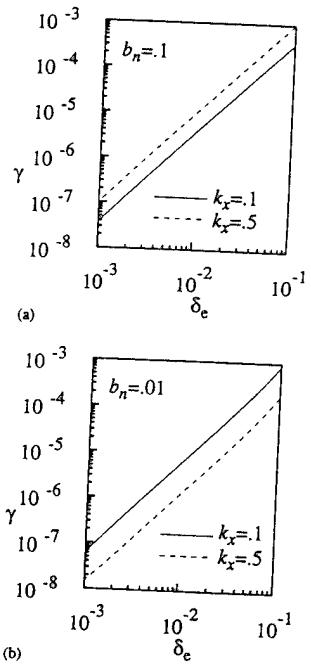

FIG. 11. Growth rate dependence on the electron inertial length for two values of the wave number k_x. (a) b_n=0.1. (b) b_n=0.01. For these cases, the growth rate is proportional to δ_e^2.

$b_n \sim 10^{-2}$, with a local maxima that shifts to larger b_n with increasing wave number. The variation of the growth rate with b_n is about an order of magnitude over five decades of b_n and is independent of b_n for $b_n < 10^{-3}$. All modes are stabilized completely for sufficiently large b_n: the k_x=0.5 mode is stabilized for $b_n > 0.35$, and the k_x=0.1 mode is stabilized for $b_n > 0.2$ (for the indicated values of δ_e). This form of electron tearing is extremely weak, particularly if $b_n \neq 0$. If the effect of plasma compressibility is included, we expect that the collisionless mode will be stabilized for significantly lower values of b_n, as demonstrated in Refs. 5 and 29.

The dependence of the collisionless tearing growth rate on the electron inertial length is illustrated in Fig. 11. For the indicated values of wave number and b_n the growth rate in this range of δ_e varies as δ_e^2. This dependence yields lower growth rates than that of Ref. 30 ($\gamma \sim k_x \delta_e$). Since we have confirmed the correctness of the result of Ref. 30 earlier for b_n=0, we attribute the different scaling to the choice of the range of k_x and δ_e, and more importantly, to the presence of a normal magnetic field component, which weakens the instability considerably.

To see the combined effect of the resistive and electron inertial terms in Ohm's law, we note that these effects enter the equations in the form (ignoring the spatial effect of the normalized density) $S^{-1} - i\omega\delta_e^2$; so for electron inertia to have a significant effect on resistive tearing, it is necessary to have $|\omega|\delta_e^2 S \geq 1$. In the presence of b_n, since $|\omega|$ is small and δ_e is a very small number, one has to consider extremely large values of S before the effect of δ_e can be visible.

We emphasize that the simple form (1) of the generalized Ohm's law does not include all of the relevant collisionless physics of the magnetotail. For instance, in the absence of resistivity, the electron inertia term produces narrow nonideal spatial scales of the order of δ_e that are likely to be altered qualitatively by the presence of other terms (such as the pressure tensor term). Such a simple Ohm's law can treat a restricted form of electron tearing (as in Ref. 29). Despite the simplicity of our Ohm's law, we have shown here that the normal magnetic field has a strongly stabilizing effect on the magnetotail, independent of whether a collisional or a collisionless mechanism breaks field lines.

V. DISCUSSION AND SUMMARY

The nominal near-earth magnetotail plasma sheet is approximately parametrized by $b_n \sim 10^{-1}$, $\delta_e \sim 10^{-3}$ (in units of B_0 and current sheet half-thickness a) and an Alfvén time $\tau_A \sim 10$ s, which normalizes our numerical results. The Lundquist number for a resistivity based on classical collisions is typically $S \sim 10^{12}$. It has been argued that microturbulence can cause enhancements of η by several orders of magnitude. Even for S values as small as 10^4, Fig. 6 suggests that for the nominal magnetotail, the growth rate of the resistive tearing mode is of the order $\gamma \sim 10^{-4}$ s. This is a very weak instability, and cannot be a candidate, even for the sluggish growth phase of a substorm, which can last as long as 1.5 h.

Attention then shifts to the collisionless tearing instability. The growth rate of the type of electron tearing considered in this paper also decreases rapidly at values of a B_n characteristic of the nominal magnetotail. We believe, as discussed in Refs. 17 and 24, as well as in the Introduction of this

paper, that the stabilization found analytically as well as numerically here is symptomatic of a deeper cause that has to do with geometry. Unless $B_n=0$, there can be no genuine topological change due to a linear instability, and the pictures presented in the Introduction truly capture the essence of many elaborate calculations that find stability with respect to tearing modes in the presence of B_n. The bunching of neighboring field lines in Fig. 2 suggests that plasma compressibility is likely to play a strong stabilizing role. It should be noted that we have chosen to omit the compressibility effects here at the very outset due to our choice of the representation of $\boldsymbol{\nu}$, and the modes stabilize due to the effect of B_n. If compressibility is included, as in Refs. 5 and 24, the stabilizing role of B_n would be seen to be even more drastic. These conclusions do not change significantly if a constant B_y field is included in the analysis, unless B_n itself is very small.

In the last 20 years of research on collisionless tearing modes in the magnetotail, of which Refs. 1–10 is a partial record, there has been considerable emphasis on sophisticated calculations of particle orbits and the possible ways in which the details of orbits may alter the complex conductivity tensor that relates the perturbed current density and the perturbed electric field. Too little emphasis has been given to considerations of geometry and topological change, which, in our present view,[17] is at the heart of the debate on reconnection if it is to be distinguished from simple diffusion. Geometric considerations in the presence of B_n suggest to us that the tearing instability (collisional or collisionless) cannot be found to grow as a robust instability merely by adding additional refinements to Ohm's law (such as pressure anisotropy[31]) unless the equilibrium (2) is qualitatively altered so that B_n is reduced, at least locally, to near-zero values. Such a qualitative alteration of the equilibrium may be affected in a number of ways, discussed, for instance in Refs. 17 and 24. Unless such qualitative changes occur, the plausibility of tearing instabilities as a mechanism for substorm onset at near-Earth distances remain in serious doubt. However, other MHD or kinetic instabilities are possible, and are currently the subject of active research.

ACKNOWLEDGMENTS

This research is supported by Air Force Office of Scientific Research Grant No. F49620-93-1-0071 and National Science Foundation Grant No. NSF-ATM-9209006.

APPENDIX: OSCILLATORY BEHAVIOR OF PERTURBED STREAMFUNCTION

Far from the current sheet, ρ goes to zero exponentially. Taking the limit as S^{-1} and ρ go to zero in (6) and (7):

$$\left(1-\frac{b_n^2}{\omega^2\delta_e^2}\right)\frac{\partial^2\phi}{\partial z^2}$$
$$=\left(k_x^2+\frac{4ik_xb_n}{\omega^2}-\frac{k_x^2}{\omega^2\delta_e^2}-\frac{2ik_xb_n}{\omega^2\delta_e^2}\right)\phi$$
$$+2\left(-\frac{ib_nk_x}{\omega^2}+1-\frac{b_n^2}{\omega^2\delta_e^2}+\frac{ik_xb_n}{\omega^2\delta_e^2}\right)\frac{\partial\phi}{\partial z}$$
$$+\frac{ib_n}{\omega\delta_e^2}\frac{\partial A}{\partial z}-\frac{k_x}{\omega}\left(2+\delta_e^{-2}+\frac{2ib_n}{k_x\delta_e^2}\right)A, \tag{A1}$$

$$\frac{\partial^2A}{\partial z^2}=k_x^2A. \tag{A2}$$

Equation (A2) implies an exponential variation of the magnetic potential with scale length k_x^{-1}, as indicated in Fig. 9(a). Substituting (A2) into (A1) yields

$$\left(\frac{\partial^2}{\partial z^2}-k_x^2\right)\left(1-\frac{b_n^2}{\omega^2\delta_e^2}\right)\frac{\partial^2\phi}{\partial z^2}$$
$$=\left(\frac{\partial^2}{\partial z^2}-k_x^2\right)\left(k_x^2+\frac{4ik_xb_n}{\omega^2}-\frac{k_x^2}{\omega^2\delta_e^2}-\frac{2ik_xb_n}{\omega^2\delta_e^2}\right)\phi$$
$$+2\left(\frac{\partial^2}{\partial z^2}-k_x^2\right)\left(-\frac{ib_nk_x}{\omega^2}+1-\frac{b_n^2}{\omega^2\delta_e^2}+\frac{ik_xb_n}{\omega^2\delta_e^2}\right)\frac{\partial\phi}{\partial z}. \tag{A3}$$

Assuming that $\phi\sim\exp(\alpha_z z)$, we obtain $\alpha_z=\pm k_x$ and

$$(\omega^2\delta_e^2-b_n^2)\alpha_z^2-(k_x^2\omega^2\delta_e^2+4ik_xb_n\delta_e^2-k_x^2-2ik_xb_n)$$
$$-2(-ib_nk_x\delta_e^2+\omega^2\delta_e^2-b_n^2+ik_xb_n)\alpha_z=0. \tag{A4}$$

Hence, with $|\omega^2\delta_e^2/b_n^2|\ll 1$ and $|\alpha^2\omega^2\delta_e^2/b_n^4|<1$, the root with the upper sign becomes

$$\alpha_z\sim -ik_x/b_n+(1-\sqrt{1-k_x^2\omega^2\delta_e^2/b_n^2}). \tag{A5}$$

The oscillatory character in Figs. 9(b) and 9(c) is evidence that the mode with α_z scaling as Eq. (A5) appears in the eigenmode for ϕ.

[1]B. Coppi, G. Laval, and R. Pellat, Phys. Rev. Lett. **16**, 1207 (1966).
[2]K. Schindler, J. Geophys. Res. **79**, 2803 (1974).
[3]A. A. Galeev and L. M. Zelenyi, Sov. Phys. JETP **43**, 1113 (1976).
[4]F. V. Coroniti, J. Geophys. Res. **85**, 6719 (1980).
[5]B. Lembege and R. Pellat, Phys. Fluids **25**, 1995 (1982).
[6]J. Büchner and L. M. Zelenyi, J. Geophys. Res. **92**, 13456 (1987).
[7]X. Wang, A. Bhattacharjee, and A. T. Y. Lui, J. Geophys. Res. **95**, 15047 (1990).
[8]R. Pellat, F. V. Coroniti, and P. L. Pritchett, Geophys. Res. Lett. **18**, 143 (1991).
[9]M. M. Kuznetsova and L. M. Zelenyi, Geophys. Res. Lett. **18**, 1825 (1991).
[10]J. Hernandez, W. Horton, and T. Tajima, J. Geophys. Res. **98**, 5893 (1993).
[11]T. Terasawa, J. Geophys. Res. **86**, 9007 (1981).
[12]J. E. M. Hamilton and J. W. Eastwood, Plasma Phys. **39**, 293 (1982).
[13]J. Ambrosiano, L. C. Lee, and D. W. Swift, J. Geophys. Res. **88**, 7860 (1983).
[14]D. W. Swift and C. Allen, J. Geophys. Res. **92**, 10015 (1987).
[15]W. Zwingmann, J. Wallace, K. Schindler, and J. Birn, J. Geophys. Res. **95**, 20877 (1990).
[16]P. L. Pritchett, F. V. Coroniti, R. Pellat, and H. Karimabadi, J. Geophys. Res. **96**, 11523 (1991).
[17]A. Bhattacharjee, X. Wang, and M. Tanaka, in *Substorms 2*, Proceedings

of the 2nd International Conference on Substorms, Fairbanks, Alaska, 7–11 March 1994, edited by J. R. Kan, J. D. Craven, and S-I. Akasofu (University of Alaska Press, Fairbanks, AK, 1994), pp. 121–126.
[18]J. N. Leboeuf, T. Tajima, C. F. Kennel, and J. M. Dawson, Geophys. Res. Lett. **8**, 257 (1981).
[19]J. G. Lyon, S. H. Brecht, J. D. Huba, J. A. Fedder, and P. J. Palmadesso, Phys. Rev. Lett. **46**, 1038 (1981).
[20]T. Ogino, J. Geophys. Res. **91**, 6791 (1986).
[21]K. Watanabe and T. Sato, J. Geophys. Res. **95**, 75 (1990).
[22]R. J. Walker, T. Ogino, J. Raeder, and M. Ashour-Abdalla, J. Geophys. Res. **98**, 17235 (1993).
[23]H. P. Furth, J. Killeen, and M. N. Rosenbluth, Phys. Fluids **6**, 459 (196
[24]X. Wang and A. Bhattacharjee, J. Geophys. Res. **98**, 19419 (1993).
[25]C. de Boor, SIAM J. Num. Anal. **14**, 491 (1977).
[26]K-I. Nishikawa and J. Sakai, Phys. Fluids **25**, 1384 (1982).
[27]B. Coppi, R. Galvão, R. Pellat, M. N. Rosenbluth, and P. Rutherford, So J. Plasma Phys. **2**, 533 (1976).
[28]P. Rosenau, Phys. Fluids **26**, 2578 (1983).
[29]R. Iacono, A. Bhattacherjee, C. Ronchi, J. M. Greene, and M. H. Hughe Phys. Plasmas **1**, 2645 (1994).
[30]M. Ottaviani and F. Porcelli, Phys. Rev. Lett. **71**, 3802 (1993).
[31]J. Chen and P. J. Palmadesso, Phys. Fluids **27**, 1198 (1986).

GEOPHYSICAL RESEARCH LETTERS, VOL. 22, NO. 21, PAGES 2985-2988, NOVEMBER 1, 1995

Growth, sudden enhancement, and relaxation of current sheets in the magnetotail: Two-dimensional substorm dynamics

Z. W. Ma, Xiaogang Wang, and A. Bhattacharjee

Department of Physics and Astronomy, University of Iowa, Iowa City

Abstract. A two-dimensional, low-dissipation magnetohydrodynamic simulation is used as a model for magnetospheric substorm dynamics. The simulation begins from a Grad-Shafranov equilibrium of the magnetotail including the Earth's dipole field. This initial configuration is driven by a continuous electric field (of IMF origin) which induces equatorward flows. A thin current sheet develops algebraically in time, spanning Y-points that stretch from the mid-tail region ($\sim 30\,R_E$) to the near-Earth region ($\sim 10\,R_E$). In the late growth phase, rapid thinning of the sheet and near-explosive intensification of the current density are observed. The onset of a partial current disruption results from the formation of small islands in the near-Earth region. Numerical results are compared with observations.

Introduction

There is increasing appreciation that the key to the problem of substorm onset lies in the growth phase when the magnetotail is prepared for the violent relaxation dynamics that follows. Typical phenomena observed during the growth phase [*McPherron,* 1970] indicate an intensification of the cross-tail current which results in a tail-like reconfiguration of the nightside magnetic field at near-Earth distances ($< 10\,R_E$) [*Kaufmann,* 1987]. Observations show that the thickness of the current sheet can be reduced to less than $1\,R_E$ in the near-Earth region before the onset of the expansion phase [*Mitchell et al.,* 1990; *Sergeev et al.,* 1990; *Lui et al.,* 1992]. In particular, *Ohtani et al.* [1992] report a near-explosive intensification in the cross-tail current density during a short interval (< 1 min) just before onset, after a period of sluggish growth ($\sim 0.5-1.5$ hr). These observations are of great interest because they show the presence of two distinct time scales in the growth phase.

The onset of the expansion phase involves a sudden reduction of the cross-tail current in the near-Earth region ($7-11\,R_E$) [*Lui et al.,* 1992]. The current sheet disruption region is localized ($< 1\,R_E$) and the disruption of the current is also partial, typically involving about 20% of the cross-tail current [*Lui,* 1978; *Ohtani et al.,* 1992].

There have been some magnetohydrodynamic (MHD) simulations of current sheet formation and rapid reconnection in the near-Earth magnetotail driven by an external electric field that induces an equatorward boundary flow [*Lee et al.,* 1985; *Birn and Hesse,* 1991]. A common feature of all these simulations is that the characteristic Lundquist number S is typically below 10^3. An important consequence of the large value of resistivity used in these simulations is the prompt creation of a near-Earth neutral line (NENL) and rapid reconnection, with large earthward and tailward flows near the X-point, and subsequent ejection of a plasmoid, qualitatively consistent with the NENL model of substorms [*Hones,* 1979].

Even if we assume that anomalous resistivity is indeed generated by some form of fluid or microturbulence, it is not obvious whether one can rely on such turbulence to bridge the enormous gap between the classical Spitzer value and the values used in the numerical studies cited above. Then an interesting issue that deserves attention is the dynamics of the magnetotail in the high-S regime, and the qualitative differences with the low-S regime.

In the present paper, we present numerical results from a two-dimensional, high-S simulation of the magnetotail, driven by sub-Alfvénic equatorward boundary flows. We assume the resistivity to be constant, without an ad hoc current-dependent or spatially localized enhancement factor. We integrate numerically the compressible resistive MHD equations using a Runge-Kutta finite-differencing scheme that has an accuracy of fourth order in time and second order in space [*Ma et al.,* 1995]. Exploiting the symmetry of the initial conditions, we carry out the numerical simulation in half of the physical domain in the $x-z$ plane, with $x \in (-6, -52)R_E$ and $z \in (0, 15)R_E$. With a 151×121 array and a nonuniform mesh, we are able to resolve $0.05\,R_E$ in x and $0.03\,R_E$ in z. We too are limited by the available numerical resolution, but the Lundquist number specified in our simulations is approximately 10^5 (including the effect of numerical resistivity) which is about two orders of magnitude higher than in previous numerical studies, and high enough to explore an asymptotic regime in which several qualitatively new dynamical features become evident.

Copyright 1995 by the American Geophysical Union.

Paper number 95GL02937
0094-8534/95/95GL-02937$03.00

Numerical Model and Simulation Results

We use the standard GSM coordinates and assume that y is an ignorable coordinate. All variables are cast in dimensionless form: in particular, distances are scaled by 1 R_E, and time is scaled by the characteristic Alfvén time ($\sim$ 6 s). The magnetic field $\mathbf{B}$ is represented as

$$\mathbf{B} = \nabla\Psi \times \hat{\mathbf{y}} + B_y\hat{\mathbf{y}} \ . \qquad (1)$$

The initial equilibrium is obtained by solving the Grad-Shafranov equation numerically. We start with the magnetic field of the 1989 Tsyganeko model [*Tsyganeko*, 1989], specify two free flux functions and iterate numerically till we converge to an equilibrium. Due to the presence of a significant B_z-component at near-Earth distances, such an equilibrium is generally stable with respect to resistive tearing instabilities [*Harrold et al.*, 1995]. Since free reconnection is thus ruled out, we are motivated to investigate the role of forced reconnection, driven by the solar wind.

An electric field,

$$E_y = E_0\{0.6 + 0.4\tanh[(3(x_1 + x) + 2x_2)/x_2]\} \ , \qquad (2)$$

is imposed at $z = \pm 15\,R_E$, where $E_0 = 0.04$, $x_1 = 6R_E$, and $x_2 = 46R_E$. The functional form (2) is a simple representation of an electric field that peaks at near-Earth distances and decays monotonically with increasing $|x|$. The boundary perturbation results roughly in a uniform IMF outside of a flared magnetopause. The equatorward inflow v_z due to the electric field (2) is of the order of 40 km/s, which is less than 4% of the characteristic Alfvén speed ($\sim$ 1000 km/s), roughly consistent with observations. Free boundary conditions are used at the earthward and tailward boundaries of the simulation box for all dependent variables except Ψ which obeys $d\Psi/dt = 0$.

(a) Early Current Sheet Thinning and Y-type Geometry

The equatorward flows at the lower and upper tail boundaries are generated by the imposed electric field (2). The plasma flow is mainly earthward in the near-Earth region and tailward in the flared mid- and distant tail. An X-type neutral line with an extremely small separatrix angle is formed in the region around $x = -30\ R_E$, where the B_z-field is initially weak, and eventually stretches to form Y-points, with a thin current sheet of finite length spanning the Y-points [*Syrovatskii*, 1971], which is a generic feature of high-S simulations of forced reconnection [*Biskamp*, 1986; *Ma et al.*, 1995].

Whereas the tailward end of the thin current sheet remains approximately at $x = -30\ R_E$, the earthward end stretches as the earthward Y-point penetrates into the near-Earth region. This is evident by inspection of Figures 1a-c for the cross-tail current J_y. (Note that $J_y = J_y(\Psi)$ in the initial equilibrium.) The dipole magnetic field close to the Earth causes the Poynting flux to be directed towards the near-Earth region where it enhances the thinning of the plasma sheet. The current sheet spanning Y-points penetrates to near-Earth distances, with a length of approximately 20 R_E along the neutral sheet.

We are now well into the growth phase, and in real time, nearly 30 min ($300\tau_A$) into the simulation. The magnetic flux piles up locally in the near-Earth region due to the externally driven inflow, and reconnects in the mid-tail region. Small secondary islands are seen to form within the current channel (Fig. 1c), but these secondary islands are too small to break up the spatially extended Y-point geometry.

(b) Near-explosive Current Intensification

Before continuing, we summarize some recent analytical results [*Wang et al.*, 1995] which are useful in interpreting the results of the present simulation. *Wang et al.* [1995] show that under the influence of inward boundary flows in a Harris equilibrium, the current sheet (spaning Y-points) exhibits a transition from a sluggish linear growth phase to a rapid nonlinear phase in a characteristic time scale $\tau_N = (\tau_R\tau_0\tau_A^3)^{1/5}$, where τ_0 is the characteristic time scale of the imposed boundary flow, τ_A and τ_R are the characteristic Alfvén and resistive diffusion time scales, respectively. (Note that the new characteristic time scale τ_N has a much weaker dependence on resistivity ($\sim \eta^{-1/5}$) than the characteristic time scale in the Sweet-Parker model ($\sim \eta^{-1/2}$)). Furthermore, the current sheet amplitude at the separatrix grows as $J_0(t/\tau_A)^2$ in the linear regime, and as $(\tau_R/\tau_0)^{1/2}J_0(t/\tau_A)^{3/2}$ in the nonlinear regime, where J_0

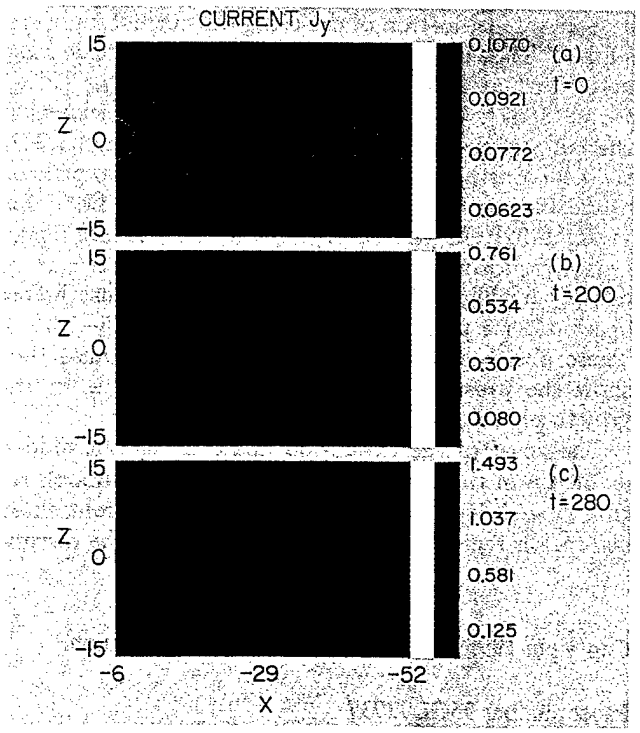

Figure 1. Projection of level surfaces of the cross-tail current on the $x - z$ plane at different times during dynamical evolution.

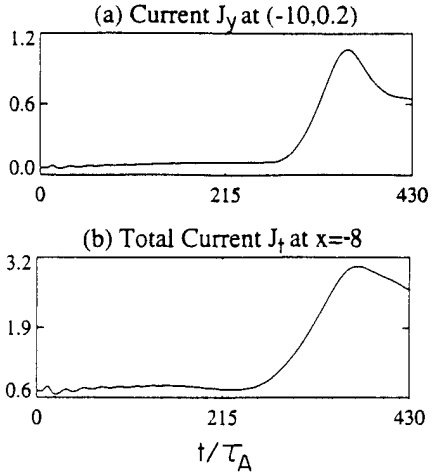

Figure 2. Time-history of (a) J_y at $x = -10$, $z = 0.2$, and (b) J_t at $x = -8$ (after integrating over z).

is a constant, independent of η. The large multiplicative factor $(\tau_R/\tau_0)^{1/2}$ accounts for the rapid enhancement of the current sheet.

The initial state (Fig. 1a) of our magnetotail simulation includes the Earth's dipole field which makes an analytical treatment difficult. However, as the plasma sheet thins and the B_z-field is reduced substantially over an extended region, the current sheet spanning Y-points penetrates to the near-Earth region (Fig. 1c), and the geometry (from $x = -10\,R_E$ tailward to about $x = -30\,R_E$) resembles qualitatively the geometry of a Harris sheet. In Figure 2a, we plot the cross-tail current J_y as a function of time at the near-Earth point $x = -10\,R_E$, $z = 0.2\,R_E$. We note a rapid enhancement in the amplitude of the current density at $t = 280\tau_A$, qualitatively similar to that seen in our earlier Harris sheet studies. This strong enhancement of the cross-tail current density (by more than 20 times of its initial value) occurs in the thin current sheet while the lobe magnetic field remains almost unchanged. (As mentioned above, nonuniform grid points are used to resolve the thin current sheet; there are approximately 50 grid points in the thin current sheet.) By this time, the overall thickness of the plasma sheet has dropped to less than 5% of its initial width.

Figure 2 shows the time evolution of (a) the current density J_y and (b) the integrated current J_t at $x = -8\,R_E$. The sudden intensification in the cross-tail current density J_y is seen to be more than 20 times its initial local value within about 6 min (Fig. 2a). This time scale in our simulation is somewhat slower than the time scale reported by *Ohtani et al.* [1992]. The analytical scaling derived by *Wang et al.* [1995] suggests that the Lundquist number in our simulations must be 50–100 times higher than what it is if we are to realize a sudden intensification time scale of 1 min or less. Such high Lundquist numbers will yield results that are in better agreement with observations, but cannot be simulated reliably using our present code.

In phase with the explosive enhancement of the current density J_y, the magnetic field B_x and the integral current J_t also exhibit rapid temporal changes. The flow-component v_x during this entire growth phase, before the onset of the expansion phase, continues to be dominantly earthward, although it is reduced locally in the vicinity of secondary islands in the near-Earth region (Fig. 3). This is qualitatively consistent with observations at $|x| < 20\,R_E$ [*Baumjohann et al.*, 1990].

(c) Partial Current Disruption

After the near-explosive growth phase, the onset of the substorm expansion occurs in a typical event when the current density J_y decreases and the magnetic field dipolarizes in the earthward side of the near-Earth Y-point. We indeed observe that the integrated current decreases by about 20% (Fig. 2b) due to the occurrence of current disruption. In our two-dimensional simulation, this disruption coincides with the expulsion of the flux rope due to enhanced reconnection at the Earthward end of the current sheet as the near-Earth Y-point starts moving tailward. At this point, the flow velocity in the disruption region changes direction and becomes tailward. Note that the flow remains always earthward in Figure 3a. It is also earthward in Figure 3b till the occurrence of the partial current disruption. The tailward flow in our simulations is confined in a narrow layer along the z-direction with a thickness of the order of $2\,R_E$.

The simulation described above is two-dimensional (i.e., $\partial/\partial y = 0$ for all times) and cannot describe the full complexity of substorm evolution in the expansion phase. In particular, important features of the three-dimensional convection pattern, the current diversion and the closure of the current system in the ionosphere cannot be described by a 2D model. By constraining the dynamics to be 2D, it is as if we have not left the magnetotail much room to evolve except to undergo reconnection as a mechanism for partial current disruption. This is the main reason why we do not dwell further on the results of the simulation after the onset of the expansion phase. It is quite possible that the stressed magnetotail with the embedded thin current sheet may be unstable to instabilities that break the symmetry condition $\partial/\partial y = 0$. Such instabilities may have either an ideal MHD character (such as a current-driven kink or a pressure-driven ballooning considered by *Roux et al.* [1991]) or be of kinetic origin (such as the cross-field current instability considered by *Lui et al.* [1991]). If these instabilities can be shown to exist with appropriate ionospheric boundary conditions, they can poten-

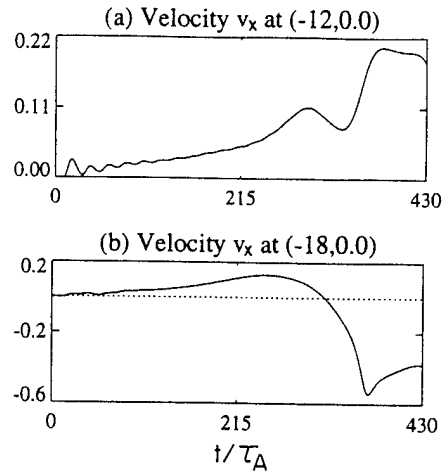

Figure 3. Time-history of the flow velocity v_x.

tially accelerate the transition time to the dipolar state which is much slower in our two-dimensional simulation than observations suggest.

Summary and Discussion

In summary, we describe the following possible scenario for a substorm. At first, under the influence of steady inward boundary flows attributable to the solar wind, an X-type line with a very small separatrix angle is formed in the mid-tail region. The reason why the mid-tail is favored for initial separatrix formation over the near-Earth region is because the B_z field is weaker at mid-tail. The mid-tail is also favored over the distant tail because the driving electric field (2) is stronger in the mid-tail than in the distant tail. If the Lundquist number is high, this short-lived X-type geometry is immediately transformed to a geometry with a thin current sheet spanning Y-points that stretches from the mid-tail to the near-Earth region.

Following a sluggish growth phase, forced (or driven) reconnection in this Y-type geometry leads to a near-explosive intensification of the current density in the near-Earth region (though the reconnected flux is small for high-S) just before the expansion onset.

After the sudden intensification of the cross-tail current, and further thinning induced by flux pile-up, forced reconnection in the near Earth region will cause disruption of the cross-tail current. This disruption in the present simulation diverts about 20% of the integrated current. The disturbance then propagates tailward, and a flux rope (with $B_y \neq 0$) is expelled as the magnetic field dipolarizes in the near-Earth region. We have discussed at the end of the second section that the real dynamics of the magnetotail at the expansion onset are likely to involve three-dimensional features not included in our two-dimensional simulation.

We conclude with comparisons to two related major substorm models: the NENL model and the current-disruption (CD) model. A key observational difficulty of the NENL model is the lack of strong tailward flow in the near-Earth region which should ostensibly be a consequence of near-Earth X-line reconnection. A modified X-line model, which relocates the X-line to the mid-tail beyond 20 R_E downstream, cannot be easily reconciled with the so-called "Kiruna conjecture" [*Kennel*, 1992] that places the location of the expansion onset at near-Earth distances. Our present results suggest that these difficulties may be resolved by the concept of Y-points that extend from the near-Earth region to the mid-tail region connected by a thin current sheet (of spatially nonuniform amplitude). In this geometry, we do not obtain the strong tailward flow and the large extended region of a southward B_z-field that are characteristics of the NENL model. Since the tailward Y-point is located at mid-tail, there can be tailward convection in the mid-tail region, whereas the earthward Y-point in the near-Earth region can witness sudden current intensification in the last stages of the growth phase. The internal stresses in the near-Earth magnetotail due to the occurrence of this near-singular current sheet superposed on a dipolar field, drives the current sheet to nonequilibrium, a manifestation of which is the current disruption process envisioned by the CD model.

Acknowledgments. We thank C. Y. Huang for helpful discussions. This research is supported by the Air Force Office of Scientific Research grant F49620-93-1-0071 and by the National Science Foundation grant ATM 93-96280.

References

Baumjohann, W., et al., Characteristics of high-speed ion flows in the plasma sheet, *J. Geophys. Res., 95,* 3801, 1990.

Birn, J., and M. Hesse, The substorm current wedge and field-aligned currents in MHD simulations of magnetotail reconnection, *J. Geophys. Res., 96,* 1611, 1991.

Biskamp, D., Magnetic reconnection via current sheets, *Phys. Fluids, 29,* 1520, 1986.

Harrold, B. G., et al., Tearing stability of the two-dimensional magnetotail, *Phys. Plasmas,* in press, 1995.

Hones, E. W., Jr., Transient phenomena in the magnetotail and their relation to substorm, *Space Sci. Rev., 23,* 393, 1979.

Kaufmann, R. L., Substorm currents: Growth phase and onset, *J. Geophys. Res., 92,* 7471, 1987.

Kennel, C. F., The Kiruna conjecture: The strong version, Kiruna, Sweden, 23-27 March 1992, *Proc. First Intl. Conf. Substorms* (ICS-1), p. 599, ESA SP-335, May 1992.

Lee, L. C., et al., A simulation study of forced reconnection processes and magnetospheric storms and substorm, *J. Geophys. Res., 90,* 10,896, 1985.

A. T. Y., Estimates of current changes in the geomagnetotail associated with a substorm, *Geophys. Res. Lett., 5,* 853, 1978.

Lui, A. T. Y. et al., A cross-field current instability for substorm expansion, *J. Geophys. Res., 96,* 11,389, 1991.

Lui, A. T. Y., et al. , Current disruptions in the near-Earth neutral sheet region, *J. Geophys. Res., 97,* 1461, 1992.

Ma, Z. W., et al., Dynamics of current sheet formation and reconnection in two-dimensional coronal loops, *Phys. Plasmas, 2,* 3184, 1995.

McPherron, R. L., Growth phase of magnetospheric substorm, *J. Geophys. Res., 75,* 5592, 1970.

Mitchell, D. G., et al., Current carriers in the near-Earth cross-tail current sheet during substorm growth phase, *Geophys. Res. Lett., 17,* 583, 1990.

Ohtani, S., et al., Initial signatures of magnetic field and energetic particle fluxes at tail reconfiguration: Explosive growth phase, *J. Geophys. Res., 97,* 19,311, 1992

Roux, A., et al., Plasma sheet instability related to the westward traveling surge, *J. Geophys. Res., 96,* 17,697, 1991.

Sergeev, V. A., et al., Current sheet thickness in the near-Earth plasma sheet during substorm growth phase, *J. Geophys. Res., 95,* 3819, 1990.

Syrovatskii, S. I., Formation of current sheets in a plasma with a frozen-in strong magnetic field, *Sov. Phys. JETP, 33,* 933, 1971.

Tsyganenko, N. A., Magnetospheric magnetic field model with a warped tail current sheet, *Planet. Space Sci., 37,* 5, 1989.

Wang, X., et al., Fast reconnection and sudden enhancement of current sheets due to inward boundary flows, submitted for the Special Issue: Invited and Review Papers from the 37th Annual Meeting of the Division of Plasma Physics of the American Physical Society, 6-10 Novemeber 1995, *Phys. Plasmas.*

A. Bhattacharjee, Z. W. Ma, and Xiaogang Wang, Department of Physics and Astronomy, University of Iowa, Iowa City, IA 52242.

(received June 12, 1995; revised August 14, 1995; accepted September 8, 1995.)

Fast magnetic reconnection and sudden enhancement of current sheets due to inward boundary flows*

Xiaogang Wang,[†,a)] Z. W. Ma, and A. Bhattacharjee
Department of Physics and Astronomy, The University of Iowa, Iowa City, Iowa 52242

(Received 6 November 1995; accepted 26 January 1996)

Magnetic reconnection is widely believed to be involved in dynamical phenomena such as solar flares or magnetospheric substorms. The Sweet–Parker model of magnetic reconnection in a Y-type geometry predicts a characteristic time scale proportional to $S^{1/2}$ (where S is the Lundquist number), which is too slow to account for the observed time scales. The Petschek model, in contrast, predicts a time scale proportional to $\ln S$ in an X-point geometry. Numerical magnetohydrodynamic (MHD) simulations in the high-S regime generally validate the Sweet–Parker model, unless the resistivity is enhanced in the diffusion region to large values (such that typically $S<10^3$). It is demonstrated in this paper that nonlinear reconnection dynamics in a Harris sheet driven by inward boundary flows occurs on a nonlinear time scale that is proportional to $S^{1/5}$ and thus has a weaker dependence on resistivity than the Sweet–Parker time scale. The current sheet amplitude at the separatrix (spanning Y points) grows algebraically in the linear regime but is suddenly enhanced after it makes a transition to the nonlinear regime. An analytical calculation is given for both the linear and the nonlinear regimes, and supported by two-dimensional resistive MHD simulations. The features of sudden current sheet enhancement and fast reconnection, controlled by boundary flows, are relevant to the phenomena of substorm onset or the impulsive phase of flares. © *1996 American Institute of Physics.* [S1070-664X(96)93505-0]

I. INTRODUCTION

Forced magnetic reconnection is often cited as a possible cause of dynamical space plasma phenomena such as solar flares, flux transfer events, or substorms. It provides a mechanism by which magnetic fields embedded in plasmas can undergo topological rearrangement due to perturbations imposed at the boundary, and thus liberate free energy that can be converted to thermal and kinetic energy. As to whether magnetic reconnection can account for the complex observational signatures of substorms or flares remains an open question, but it is clear that any theory of reconnection that proposes to account for the complexity of these phenomena must reproduce at least two significant observational features. The first feature is a time scale fast enough to account for the observed temporal signatures, and the second, more subtle, feature is the time development of the event, which jumps from a sluggish growth phase to a fast, impulsive phase. Steady-state reconnection models such as the classical ones due to Sweet–Parker[1,2] and Petschek[3] attempt to obtain the first feature, but since the dynamics is assumed to be steady state, these models are precluded from describing the second feature. An example of the second feature is the time history of growth of current sheets in the magnetotail during the growth phase of a substorm, prior to the onset of the expansion phase. Observations indicate two distinct phases in the growth of current sheets: a period of slow growth lasting approximately an hour, followed by a short period of sudden intensification lasting less than a minute.[4,5]

Sweet and Parker[1,2] showed that in a two-dimensional (2-D) magnetic field, such as $\mathbf{B}=\hat{\mathbf{y}}B_0 \tanh x/a$, with the field pointing in opposite directions on the two sides of a neutral line at $x=0$, steady-state reconnection occurs on a characteristic time scale $\tau_{SP}\equiv(\tau_A\tau_R)^{1/2}$, where $\tau_A\equiv a/v_A = a(4\pi\rho)^{1/2}/B_0$ is the Alfvén time, and $\tau_R\equiv 4\pi a^2/\eta c^2$ is the resistive diffusion time. (Here ρ is the mass density, η is the resistivity of the plasma, and c is the speed of light.) However, although τ_{SP} is much shorter than the diffusion time scale τ_R, it is still much too long to account for fast events such as flares. In order to resolve this difficulty, Petschek proposed another steady-state model with an X-point geometry.[3] In contrast with the Sweet–Parker model, Petschek's model yields reconnection rates that are near-Alfvénic, with a weak logarithmic dependence on η.

The dynamical realizability of the Petschek model has been the subject of intense controversy and research over the last three decades. (See, for instance, the interesting review by Biskamp in his recent monograph.[6]) One point of view,[6–10] supported by numerical simulations, is that the Petschek model does not include a proper treatment of the diffusion region for large values of the Lundquist number $S(\equiv\tau_R/\tau_A)$, and unless this problem is overcome by a substantial local enhancement of the resistivity near the X point, which reduces the length of the diffusion region, it is not possible to sustain a Petschek-like geometry.[6,8,10] As a result, the dynamical evolution of a system, characterized by low values of η, leads inevitably to a geometry with current sheets between Y points, as proposed by Syrovatsky,[11] with scaling properties characteristic of the Sweet–Parker model. Another point of view argues that fast reconnection rates of the type envisioned by Petschek are indeed realizable by a suitable design of boundary conditions. (See, for instance, the review by Forbes and Priest[12] for a discussion of

*Paper 8IA1, Bull. Am. Phys. Soc. **40**, 1827 (1995).
[†]Invited speaker.
[a)]Electronic mail: xwang@iowa.physics.uiowa.edu

Petschek's model and its variants.) This point of view is also supported by simulations.[8,10,12,13–16]

Upon first glance, the two points of view discussed above may appear to be in conflict, but in point of fact, they are complementary once we realize that the key parameter that separates the two different types of numerical results is the value of S, the Lundquist number. The published numerical results supporting Petschek-like reconnection scenarios with X points tend to use Lundquist numbers of the order of 10^3 or less,[8,10,12,13–16] whereas current sheets spanning Y points in a Sweet–Parker geometry typically use Lundquist numbers at least as large as 10^4. The question as to which regime of S is physically relevant does not have a clear answer, but in the low-S simulations, appeal is generally made to some form of micro- or small-scale fluid turbulence. If, on the other hand, we take the point of view that such large, orders-of-magnitude enhancement of the resistivity is an *ad hoc* assumption and implausible, then it appears that the classical steady-state models of Sweet–Parker[1,2] and Petschek[3] leave us with a quandary for high-S plasmas, even with regard to the first observational feature mentioned above. Whereas the Sweet–Parker time scale is realized dynamically in some models of forced reconnection, it is too slow. On the other hand, the Petschek model, which yields a faster time scale, appears not to be realizable in the high-S regime.

In this paper, we give new analytical as well as numerical results on a model of forced (or driven) reconnection with an inward boundary flow. This model is stimulated by the earlier work of Sato and Hayashi,[13] who claim that a locally enhanced, current-dependent resistivity and plasma compressibility play crucial roles in their simulation.[16] Our model differs significantly from their model, in that the plasma is assumed to be incompressible and in a high-S ($\geqslant 10^4$) regime. Furthermore, the resistivity is assumed to be spatially uniform, with no *ad hoc* enhancement in the current sheet.

Two new results are obtained and supported by asymptotic analysis as well as simulation. First, we demonstrate that the reconnection in the nonlinear regime occurs on a fast time scale that depends weakly on the resistivity, in contrast with the stronger dependency in the Sweet–Parker model. Yet this model conforms essentially to the geometry envisioned by Sweet, Parker, and Syrovatsky, with current sheets spanning Y points, and not to the X-point geometry of Petschek. Second, we show that the current sheet amplitude (as well as the electric field produced by reconnection) exhibits a sudden transition from a sluggish growth phase to an impulsive phase. This latter feature has an important application to magnetosphere substorm signatures,[4,5] but as best as we know, has not been reported in any previous analytical or numerical studies of forced reconnection in the high-S regime.

Our initial conditions and main results are as follows. We assume that the initial magnetic field, which has the form $\mathbf{B}=\hat{\mathbf{y}}B_0x/a$ near $x=0$, is bounded by two boundaries, $x=\pm a$. At $t=0$, steady inward flows of the form $V_0(1+\cos ky)$ are imposed at $x=\pm a$. (We choose the inward flows as steady, motivated by observations of the substorm growth phase driven by a solar wind electric field, which varies slowly with time.) We show that the subsequent dynamics in the high-S regime occurs on a sequence of characteristic time scales. In the linear phase of the dynamics (Sec. I), characterized by the time scale $\tau_L \equiv \tau_A^{2/3}\tau_R^{1/3} \ll \tau_{SP}$, a current sheet forms at the separatrix with an amplitude growing algebraically as $J_0(t/\tau_A)^2$, where J_0 is a constant independent of η. This short-lived linear phase is followed by a nonlinear phase (Sec. II) in which the reconnection scales as $(t/\tau_N)^{3/2}$, where $\tau_N \equiv (\tau_R\tau_0\tau_A^3)^{1/5}$ is a new nonlinear time scale, with $\tau_A \ll \tau_0 \equiv a/V_0 \ll \tau_{SP}$. Compared with the Sweet–Parker time scale τ_{SP} (which is proportional to $\eta^{-1/2}$), the new nonlinear time scale τ_N has a weaker dependence on η (proportional to $\eta^{-1/5}$). The growth of the current sheet amplitude exhibits a jump at the transition from the linear to the nonlinear regime. The current sheet amplitude grows as $J_0(t/\tau_A)^2$ in the linear regime, and then as $(\tau_R/\tau_0)^{1/2}J_0(t/\tau_A)^{3/2}$ in the nonlinear regime. The onset of the nonlinear regime of current sheet growth at $t<\tau_N$ is sharp, primarily due to the presence of the multiplicative factor $(\tau_R/\tau_0)^{1/2}$, which can be very large in the high-S regime. These analytical results are supported by our numerical simulations, discussed in Sec. III.

II. THE LINEAR PHASE

We consider the linear phase of the dynamics, derived from an initial-value calculation. We follow the methodology of Hahm and Kulsrud.[17] The magnetic and velocity fields are represented as

$$\mathbf{B}(x,y,t)=\hat{\mathbf{z}}\times\nabla\psi(x,y,t) \tag{1a}$$

and

$$\mathbf{v}(x,y,t)=\hat{\mathbf{z}}\times\nabla\phi(x,y,t), \tag{1b}$$

respectively. For the initial equilibrium,

$$\mathbf{B}_0=\hat{y}B_0\tanh x/a+B_0\hat{\mathbf{z}}\cong\hat{\mathbf{y}}B_0x/a+B_0\hat{\mathbf{z}}, \quad |x|\leqslant a, \tag{2}$$

the flux and streamfunctions are given, respectively by $\psi_0=B_0x^2/2a$ and $\phi_0=0$. Inward boundary flows of the form $V_0(1+\cos ky)$, imposed at $x=\pm a$, perturb the equilibrium boundary. The time-dependent Lagrangian boundary condition for a fluid element is

$$x=\pm[a-V_0t(1+\cos ky)]. \tag{3}$$

We write $\psi(x,y,t)=\psi_0(x)+\tilde{\psi}(x,y,t)$. In the linear phase, the perturbed flux function $\tilde{\psi}(x,t)$ can be written as

$$\tilde{\psi}(x,y,t)=\tilde{\psi}_0(x,t)+\tilde{\psi}_k(x,y,t)\cos ky. \tag{4}$$

The linearized resistive MHD equations then yield

$$\frac{\partial\tilde{\psi}}{\partial t}+\nu_x\frac{B_0x}{a}=\frac{a^2}{\tau_R}\left(\frac{\partial^2}{\partial x^2}-k^2\right)\tilde{\psi}, \tag{5a}$$

for Faraday's equation, and

$$4\pi\rho\frac{\partial}{\partial t}\left(\frac{\partial^2}{\partial x^2}-k^2\right)\phi=-\frac{B_0x}{a}\frac{\partial}{\partial y}\left(\frac{\partial^2}{\partial x^2}-k^2\right)\tilde{\psi}, \tag{5b}$$

for the vorticity equation.

Since $\tau_A \ll \tau_0 \ll \tau_R$ for forced reconnection, the plasma obeys the ideal magnetostatic equations in the outer region,

whereas the effects of resistivity and inertia are important in the inner region. The outer region equations are

$$\frac{\partial \tilde{\psi}}{\partial t}+v_x\frac{\partial \psi_0}{\partial x}=0 \tag{6}$$

and

$$\frac{\partial^2 \tilde{\psi}}{\partial x^2}-k^2\tilde{\psi}=0. \tag{7}$$

The solutions for the $k=0$ and $k\neq 0$ components of $\tilde{\psi}$ in the outer region are given by

$$\tilde{\psi}_0(x,t)=\tilde{\psi}_0(0,t)+B_0\delta_0\frac{|x|}{a}\frac{t}{\tau_A} \tag{8}$$

and

$$\tilde{\psi}_k(x,t)=\tilde{\psi}_k(0,t)\left(\cosh kx-\frac{\sinh|kx|}{\tanh ka}\right)+B_0\delta_0\frac{\sinh|kx|}{\sinh ka}\frac{t}{\tau_A}, \tag{9}$$

respectively, where $\delta_0\equiv V_0\tau_A$. If the plasma is ideal, the topology of magnetic field lines cannot change, and $\tilde{\psi}(0,t)$ remains zero for all times. Then, a tangential discontinuity (current sheet) occurs in the ideal solution, with a corresponding jump in $\tilde{\psi}$.[17] The perturbed flow,

$$v_x=\mp V_0\left(1+\frac{a}{x}\frac{\sinh|kx|}{\sinh ka}\cos ky\right), \tag{10}$$

also exhibits a jump at $x=0$.

In the presence of a small but finite resistivity, $\tilde{\psi}(0,t)$ changes from its initial value, islands open up, and the current sheet is resolved. The dynamics of the inner region can be determined by solving the inner region equations near $x=0$. In the linear phase, for $t\ll\tau_L$, the Laplace transform of the kth-Fourier component,

$$\bar{\psi}_k(s)\equiv\int_0^\infty dt\, e^{-st}\tilde{\psi}_k(t),$$

can be calculated by a straightforward method discussed in Ref. 17 to obtain

$$Z\equiv\frac{d^2\bar{\psi}_k}{d\theta^2}=\frac{s\tau_L\bar{\psi}_k(0)}{[1+(ka\theta/s\tau_L)^2]},\quad \theta\equiv\left(\frac{\tau_A}{\tau_R}\right)^{-1/3}\frac{x}{a}. \tag{11}$$

It can then be shown that[17]

$$\left.\frac{d\bar{\psi}_k}{dx}\right|_{\text{inner}}\xrightarrow{x\to\infty}\frac{\pi}{2}\frac{s^2\tau_A\tau_R}{ka^2}\bar{\psi}_k(0). \tag{12}$$

On the other hand, the outer region solution (9) yields

$$\left.\frac{d\bar{\psi}_k}{dx}\right|_{\text{outer}}\xrightarrow{x\to 0}\frac{kB_0\,\delta_0}{s^2\tau_A\sinh ka}-\frac{k\bar{\psi}_k(0)}{\tanh ka}. \tag{13}$$

Matching (12) and (13), and carrying out the inverse Laplace transform, we obtain

$$\tilde{\psi}_k(0,t)=\frac{B_0\delta_0(ka)^2}{3\pi\sinh ka}\left(\frac{t}{\tau_L}\right)^3. \tag{14}$$

Equation (11) predicts that the current sheet amplitude,

$$\tilde{J}_k(0,t)=\frac{ck^2B_0\delta_0}{4\pi^2\sinh ka}\left(\frac{t}{\tau_A}\right)^2, \tag{15}$$

increases as $(t/\tau_A)^2$. Note that in the linear regime, the current sheet amplitude is independent of the resistivity. The growth of this current sheet is faster than that in Taylor's model,[17,18] in which the growth is proportional to t/τ_A, also independent of the resistivity. From Eqs. (14) and (15), in the limit $k\to 0$, we obtain $\tilde{\psi}_0(0,t)\to 0$ and $\tilde{J}_0(0,t)\to 0$.

We repeat for emphasis that the growth of the current sheet due to forced reconnection is *algebraic in time, even in the linear regime*. In this respect, forced reconnection is very different from a linear instability such as the kink-tearing mode,[19] which grows exponentially in time in the linear regime. In linear forced reconnection, the current sheet amplitude grows algebraically in time out of the *stable* magnetohydrodynamic (MHD) continuum. This distinction between free and forced reconnection is not considered in the linear theory of Horton and Tajima,[20] who conclude that the reconnected flux in forced reconnection grows exponentially in time, with a growth rate similar to that of the kink-tearing instability.

III. THE NONLINEAR PHASE

We now consider the nonlinear regime. It has been shown by Wang and Bhattacharjee[18] that the nonlinear treatment of Taylor's model in Ref. 17 omits the current sheet that persists in the nonlinear regime.[21,22] The omission occurs in Ref. 17 because it is assumed that the system passes from a linear non-constant-ψ regime to a linear constant-ψ regime, and thereafter, to a nonlinear Rutherford regime.[23] However, as demonstrated in Ref. 18, the linear constant-ψ regime is superseded by the nonlinear Sweet–Parker regime because the island width becomes comparable to the reconnection layer width while the system is still in the linear non-constant-ψ regime. To see this, we estimate the width of the island in the non-constant-ψ regime. Along the separatrix, we have

$$-\tilde{\psi}(w)+\frac{B_0w^2}{2a}\cong\tilde{\psi}(0). \tag{16}$$

In the constant-ψ regime, we have $\tilde{\psi}(w)\approx\tilde{\psi}(0)$ and the island width is given by $w\cong 2[a\tilde{\psi}(0)/B_0]^{1/2}$. However, in the non-constant-ψ regime, we must use (8) and (9) to obtain the island width,

$$w\cong\delta_0\left(1+\frac{ka}{\sinh ka}\right)\frac{t}{\tau_A}+2\left(\frac{a\tilde{\psi}(0)}{B_0}\right)^{1/2}. \tag{17}$$

It can be shown that the ratio of the island width w to the linear reconnection layer width Δ_L is given by

$$\frac{w}{\Delta_L}\approx ka\left(1+\frac{ka}{\sinh ka}\right)\frac{t^2}{\tau_0\tau_A}+[3\pi^3\sinh(ka)]^{-1/2}(ka)^2\left(\frac{t}{\tau_N}\right)^{5/2}. \tag{18}$$

where $\tau_N \equiv (\tau_0 \tau_R \tau_A^3)^{1/5}$, and, as in Ref. 18, we have replaced Δ_L by an upper bound, the current channel width $\sim(2\pi/k)(\tau_A/t)$. Then, the transition from the linear non-constant-ψ regime to the nonlinear non-constant-ψ regime occurs when $\tau_A^{1/2}\tau_0^{1/2} < t < \tau_N \ll \tau_L$.

We now use a variant of the Sweet–Parker analysis for the nonlinear regime.[1,2,7,18,21,22] Near the separatrix, using the induction equation, we obtain

$$\frac{d\tilde{\psi}(0,t)}{dt} \approx \frac{a^2(\Delta B_y)}{\tau_R \Delta_N} = \frac{2a^2 B_y}{\tau_R \Delta_N} = \frac{2a}{\tau_R \Delta_N} B_0 \delta_0 \left(1 + \frac{ka}{\sinh ka}\right) \frac{t}{\tau_A}, \tag{19}$$

where Δ_N denotes the nonlinear reconnection layer width. On the other hand, from the inner limit of the outer (ideal) region solution, we obtain

$$\frac{d\tilde{\psi}(0,t)}{dt} \cong v_x B_y. \tag{20}$$

Using the conservation laws of mass and energy for an incompressible fluid (see, for instance, Refs. 18 and 22), we obtain

$$v_A (B_y/B_0)\Delta_N \cong v_x l, \tag{21}$$

where l is the length of the current sheet. From Eqs. (19)–(21), we obtain

$$\Delta_N \cong (2B_0 \tau_A l / B_y \tau_R a)^{1/2} a. \tag{22}$$

The nonlinear reconnected flux can now be determined by substituting Δ_N in (17) and integrating with respect to time. We obtain

$$\tilde{\psi}(0,t) \cong \frac{2\sqrt{2}}{5} \left(1 + \frac{ka}{\sinh ka}\right)^{3/2} \left(\frac{a}{l}\right)^{1/2} B_0 \delta_0 \left(\frac{t}{\tau_N}\right)^{5/2}. \tag{23}$$

As claimed earlier, the characteristic reconnection time scale is given by $\tau_N = (\tau_R \tau_0 \tau_A^3)^{1/5}$. (See the Appendix for a more detailed discussion on characteristic time scales.) In the physically interesting regime $\tau_0 \ll \tau_{SP}$, τ_N is a fast time scale, with a weak dependence on resistivity. For example, if we assume that $\tau_0 \sim 10^2 \tau_A$, $\tau_R \sim 10^8 \tau_A$, we obtain $\tau_N \sim \tau_0 \sim 10^2 \tau_A$.

The electric field near the neutral sheet is given by

$$\tilde{E}(0,t) \cong \frac{1}{c} \frac{\partial \tilde{\psi}(0,t)}{\partial t} \cong \sqrt{2} \left(1 + \frac{ka}{\sinh ka}\right)^{3/2} \left(\frac{a}{l}\right)^{1/2} \frac{B_0 \delta_0}{c \tau_N} \left(\frac{t}{\tau_N}\right)^{3/2}. \tag{24}$$

To determine a rough scaling, if we take $a \sim l$ and $ka \sim 1$; then for $t \sim \tau_N$, we obtain $\tilde{E}(0) \sim V_0^{6/5} \eta^{1/5}$, which implies that the reconnection electric field is determined predominantly by the driving flow (or the driving electric field) at the boundary, and only weakly by the resistivity, as found in the simulations of Sato and co-workers.[13,16]

The current density at the separatrix is given by

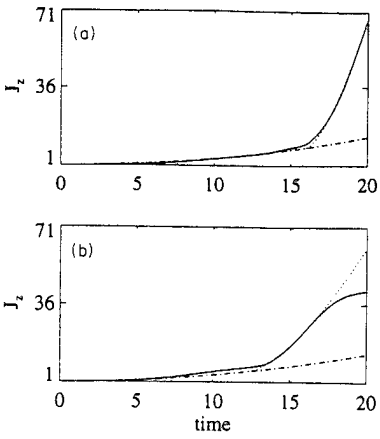

FIG. 1. The current density $J_z(0,\tau)$ as a function of (dimensionless) time $\tau(\equiv t/\tau_A)$ in (a) case (i) with $S=10^5$, and (b) case (ii) with $S=10^4$. The dash–dotted line is the linear analytical solution and the dashed line is the nonlinear one.

$$\tilde{J}(0,t) = \frac{\sqrt{2}c}{4\pi a^2} \left(1 + \frac{ka}{\sinh ka}\right)^{3/2} B_0 \delta_0 \left(\frac{a\tau_R}{l\tau_0}\right)^{1/2} \left(\frac{t}{\tau_A}\right)^{3/2}. \tag{25}$$

IV. NUMERICAL RESULTS AND DISCUSSIONS

We have tested the analytical results discussed above with a two-dimensional (2-D), incompressible MHD code. Details of the numerical method are discussed in Ref. 2. We impose inward flows at $x = \pm a$, and open boundary conditions at $y = \pm 2a$. We have run simulations with the following two sets of parameters:

(i) $\tau_0 = 100\tau_A$, $S = 10^5$, which yield $\tau_N \approx 25\tau_A$ and $\tau_L \approx 46\tau_A$;

(ii) $\tau_0 = 100\tau_A$, $S = 10^4$, which yield $\tau_N \approx 16\tau_A$ and $\tau_L \approx 21\tau_A$.

Figures 1(a) and 1(b) show the time evolution of the current sheet amplitude at $x=0$ for the sets (i) and (ii), respectively. In the linear regime, which holds in the interval $0 < t < \tau_N \ll \tau_L$, Eq. (15) predicts that the current sheet amplitude grows as $\tilde{J}_z(0,\tau) \equiv J_0 \tau^2$, where $J_0 \equiv ck^2 B_0 \delta_0 / (4\pi^2 \sinh ka)$ and $\tau \equiv t/\tau_A$. In the nonlinear regime, realized for $t > \tau_N$, we obtain

$$\tilde{J}_z(0,\tau) = \frac{\sqrt{2}\pi \sinh ka}{(ka)^2} \left(1 + \frac{ka}{\sinh ka}\right)^{3/2} \left(\frac{a\tau_R}{l\tau_0}\right)^{1/2} J_0 \tau^{3/2}. \tag{26}$$

For the computer runs reported here, we have taken $ka = \pi/2$ (for which the linear stability parameter Δ'[24] is negative). Then, Eq. (26) yields

$$\tilde{J}_z(0,\tau) \cong 9 \left(\frac{a}{l}\right)^{1/2} \left(\frac{\tau_R}{\tau_0}\right)^{1/2} J_0 \tau^{3/2}. \tag{27}$$

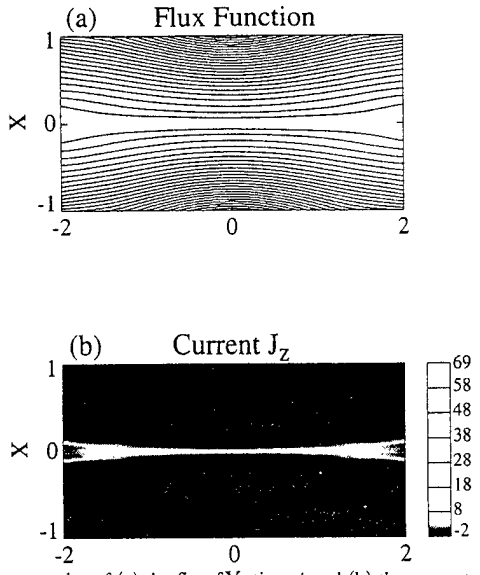

FIG. 2. Contour plot of (a) the flux function ψ and (b) the current density J_z at $\tau=20$ in case (i) with $S=10^5$.

In Eq. (27), the length of the current sheet l is determined by the wavelength of the boundary perturbation, as given by the inequality $1/k<l<\pi/k$. For the present set of computer runs, we have the inequality

$$6<9(a/l)^{1/2}<11. \tag{28}$$

In Figs. 1(a) and 1(b), we show that the simulation results (solid line) are reproduced well by the curve

$$J(\tau)=\begin{cases} J_{z0}+J_0\tau^2 & \tau<t_0 \\ J_{z0}+J_0t_0^2+\alpha\left(\dfrac{\tau_R}{\tau_0}\right)^{1/2}J_0(\tau-t_0)^{3/2} & \tau>t_0 \end{cases}, \tag{29}$$

where J_{z0} is the initial equilibrium current density at $x=0$. The dash–dotted line is the linear solution, while the dashed line is the nonlinear solution. The factor α is 7.2 for case (i) and 8.8 for case (ii). Both of these values do lie within the range of the inequality (28). The onset of the regime of enhanced growth seen to be approximately $t_0\approx16$ for case (i) and $t_0\approx12.5$ for case (ii), which should be compared with the analytical predictions of $t_0<25$ for case (i) and $t_0<16$ for case (ii). The theoretical predictions are reasonably accurate, particularly for the higher-S simulation.

Figures 2(a) and 2(b) are pictures at $\tau=20$ of the current density and the flux surface for case (i). Note the presence of an intense current sheet near the separatrix spanning the Y points. The geometry is similar to that envisioned by Sweet, Parker, and Syrovatsky, but very different from that of Petschek. For low values of $S(<10^3)$, the Y-point structure is not realized, and the geometry appears more Petschek-like.[9] We note that the flux piles up at the inflow region of the current sheet, enhancing the growth of the current sheet.

Two new features in the results discussed above become evident when we compare with earlier analytical and numerical results on forced reconnection.[7,18,21,22] In earlier work, the characteristic nonlinear reconnection time scale is τ_{SP}, which is much longer and has a stronger dependence on resistivity than the characteristic scale τ_N. The second new feature is the sudden transition in the current sheet amplitude from a slow growth phase to an impulsive phase, seen in Fig. 1(a). (Our results should also be distinguished from some recent results on fast reconnection obtained in Refs. 25–27, which begin from a highly stressed initial state containing an X point.)

These analytical and numerical results yield a possible answer to the question broached by us at the beginning of the paper: forced reconnection, if it is caused by an inward boundary flow, can indeed yield reconnection rates with a much weaker dependency on the plasma resistivity than the classical Sweet–Parker model. The reconnection electric field is determined predominantly by the magnitude of the driving electric field or the flow at the boundary. Furthermore, in the high-S regime, the growth of the current sheet amplitude exhibits a jump from the linear to the nonlinear regime. Though these are attributes of fast reconnection, they occur within the context of a Sweet–Parker geometry in a high-Lundquist-number regime, and not a Petschek geometry. These results have significant implications for phenomena that involve continuously driven reconnection such as solar coronal fields driven by foot-point convection or the Earth's magnetosphere driven by the solar wind. In particular, we have recently considered[5] the development of current sheets in the magnetotail during the growth phase of a substorm. Multisatellite observations indicate two distinct time scales: a sluggish phase that lasts an hour or so, followed by a period of sudden enhancement that lasts approximately one minute (see Fig. 13 of Ref. 4). This feature has been simulated by us,[5] using inward flows due to the solar wind in a magnetospheric equilibrium including the Earth's dipole field, where after the formation of a Y-type current sheet, reconnection forced by the solar wind leads to a sudden enhancement in the current sheet amplitude. For details the reader is referred to Ref. 5.

ACKNOWLEDGMENTS

Our interest in this problem was stimulated during a visit by one of us (A.B.) to the National Institute of Fusion Science at Nagoya during Fall, 1993, and he would like to thank Professor T. Hayashi, Professor R. Horiuchi, and Professor T. Sato for helpful discussions.

This research is supported by the Air Force Office of Scientific Research Grant No. F49620-93-1-0071 and the National Science Foundation Grant No. ATM 93-96280. Supercomputing resources were provided by the San Diego Supercomputing Center.

APPENDIX: CHARACTERISTIC TIME SCALES IN NONLINEAR RECONNECTION

The reconnection rate is defined as $\dot{\tilde{\psi}}=\partial\tilde{\psi}/\partial t$.[6] During nonsteady reconnection dynamics, this rate is, in general, a function of time. This can cause some confusion in defining a characteristic time scale in nonlinear reconnection models. A correct procedure for defining a characteristic reconnection time (in resistive MHD) is as follows: consider the explicitly time-dependent solution for the flux or the island width and

normalize the physical time by a time scale that makes the solution manifestly independent of η. That time scale defines the characteristic reconnection time.

We illustrate this definition by reviewing some well-known cases. In the Rutherford regime,[23] the island equation obeys the relation

$$w \sim \eta t, \tag{A1}$$

from which it follows that the characteristic time scale t_R scales as η^{-1}. We then also have

$$\tilde{\psi} \sim w \frac{\partial w}{\partial t} \sim \eta^2 t, \tag{A2}$$

from which we conclude, after scaling time, that the reconnection rate $\tilde{\psi}$ scales as η.

Another well-known case is nonlinear $m=1$ kink tearing.[21,22] From the island equation

$$w \sim \eta t^2, \tag{A3}$$

we obtain the characteristic time scale $\tau_{SP} \sim \eta^{-1/2}$. It also follows that

$$\frac{\partial \tilde{\psi}}{\partial t} \sim w \frac{\partial w}{\partial t} \sim \eta^2 t^3 \sim \eta^{1/2}, \tag{A4}$$

which, after time has been scaled, shows that the reconnection rate scales as $\eta^{1/2}$.

We now consider the model developed in this paper and show that by Eq. (23), the characteristic reconnection time scale is τ_N. If we write

$$\frac{\partial \tilde{\psi}}{\partial t} \sim \eta^{1/2} t^{3/2}, \tag{A5}$$

and conclude hastily that the reconnection rate scales as $\eta^{1/2}$, it will be wrong, because in Eq. (A5), time has not been scaled correctly. After t is normalized by τ_N, it follows that the reconnection rate scales as

$$\frac{\partial \tilde{\psi}}{\partial t} \sim \eta^{1/2} t^{3/2} \sim \eta^{1/5}, \tag{A6}$$

as claimed in this paper. In general, if $\tilde{\psi} \sim (\eta^\alpha t)^n$, then the characteristic reconnection time scale $t \sim \eta^{-\alpha}$, and $\tilde{\psi} = \partial \tilde{\psi} / \partial t \sim \eta^\alpha (\eta^\alpha t)^{n-1} \sim \eta^\alpha$.

[1] P. A. Sweet, in *Electromagnetic Phenomena in Cosmical Physics*, edited by B. Lehnert (Cambridge University Press, New York, 1958), pp. 123–139.
[2] E. N. Parker, J. Geophys. Res. **62**, 509 (1957).
[3] H. E. Petschek, in *AAS/NASA Symposium on the Physics of Solar Flares*, edited by W. N. Hess (National Aeronautics and Space Administration, Washington, DC, 1964), pp. 425–437.
[4] S. Ohtani, K. Takahashi, L. J. Zanetti, T. A. Potemra, and R. W. McEntire, J. Geophys. Res. **97**, 19311 (1992).
[5] Z. W. Ma, X. Wang, and A. Bhattacharjee, Geophys. Res. Lett. **22**, 2985 (1195).
[6] D. Biskamp, *Nonlinear Magnetohydrodynamics* (Cambridge University Press, Cambridge, 1993), Chap. 6.
[7] D. Biskamp, Phys. Fluids **29**, 1520 (1986).
[8] M. Scholer, J. Geophys. Res. **94**, 8805 (1989).
[9] Z. W. Ma, C. S. Ng, X. Wang, and A. Bhattacharjee, Phys. Plasmas **2**, 3184 (1995).
[10] M. Yan, L. C. Lee, and E. R. Priest, J. Geophys. Res. **98**, 7593 (1993).
[11] S. I. Syrovatsky, Sov. Phys. JETP **33**, 933 (1971).
[12] T. G. Forbes and E. R. Priest, Rev. Geophys. **25**, 1853 (1987).
[13] T. Sato and T. Hayashi, Phys. Fluids **22**, 1189 (1979).
[14] M. Ugai, Plasma Phys. Controlled Fusion **26**, 1549 (1984).
[15] E. R. Priest and T. G. Forbes, J. Geophys. Res. **97**, 16757 (1992).
[16] T. Sato, T. Hayashi, K. Watanabe, R. Horiuchi, M. Tanaka, N. Sawairi, and K. Kusano, Phys. Fluids B **4**, 450 (1992).
[17] T. S. Hahm and R. M. Kulsrud, Phys. Fluids **28**, 2412 (1985).
[18] X. Wang and A. Bhattacharjee, Phys. Fluids B **4**, 1795 (1992).
[19] B. Coppi, R. Galvao, R. Pellat, M. N. Rosenbluth, and P. H. Rutherford, Sov. J. Plasma Phys. **2**, 533 (1976).
[20] W. Horton and T. Tajima, J. Geophys. Res. **93**, 2741 (1988).
[21] W. Park, D. A. Monticello, and R. B. White, Phys. Fluids B **4**, 137 (1984).
[22] F. L. Waelbroeck, Phys. Fluids B **1**, 2372 (1989).
[23] P. H. Rutherford, Phys. Fluids **16**, 1903 (1973).
[24] H. P. Furth, J. Killeen, and M. N. Rosenbluth, Phys. Fluids **6**, 459 (1963).
[25] I. J. Craig and A. N. McClymont, Astrophys. J. **393**, 385 (1992).
[26] A. B. Hassam, Astrophys. J. **399**, 159 (1992).
[27] L. Ofman, P. J. Morrison, and R. S. Steinolfson, Astrophys. J. **417**, 748 (1993).

GEOPHYSICAL RESEARCH LETTERS, VOL. 23, NO. 13, PAGES 1673-1676, JUNE 15, 1996

Fast impulsive reconnection and current sheet intensification due to electron pressure gradients in semi-collisional plasmas

Z. W. Ma and A. Bhattacharjee

Department of Physics and Astronomy, University of Iowa, Iowa City

Abstract. A numerical simulation of forced reconnection and current sheet growth due to inward boundary flows in semi-collisional plasmas is presented, and contrasted with the results of an incompressible resistive MHD simulation in the high-Lundquist-number regime. Due to the presence of electron pressure (or Hall currents) in the generalized Ohm's law, the reconnection dynamics makes an impulsive transition from a slow linear regime to a nonlinear regime characterized by fast reconnection and current sheet intensification at a near-Alfvénic rate. The current sheet spanning Y-points in the early nonlinear regime shrinks and approaches an X-point geometry. The spatial scale of the collisionless parallel electric field is the ion skin depth, and decoupled from the spatial scale of the parallel current which is much narrower and determined by the Lundquist number.

The search for fast reconnection models has been a dominant activity in theoretical magnetospheric and solar physics research for more than three decades. Since the early sixties, two models of steady-state reconnection, both based on resistive MHD, have provided focal points for the discussion of reconnection dynamics. In the Sweet-Parker model (see, for instance, *Parker* [1979]), a two-dimensional (2D) magnetic field, such as $\mathbf{B} = \hat{\mathbf{x}} b_0 \tanh y/a$ with a neutral line at $y = 0$, undergoes steady-state reconnection on a characteristic time scale $\tau_{SP} \equiv (\tau_A \tau_R)^{1/2}$, where $\tau_A \equiv a/v_A = a(4\pi\rho)^{1/2}/b_0$ is the Alfvén time, and $\tau_R \equiv 4\pi a^2/\eta c^2$ is the resistive diffusion time. (Here ρ is the mass density, η is the resistivity of the plasma, and c is the speed of light.) In contrast with the Sweet-Parker model, Petschek's model [*Petschek,* 1964] yields reconnection rates that are near-Alfvénic, with a weak logarithmic dependence on η.

Solar flares or magnetotail substorms, phenomena to which reconnection models are often applied, exhibit considerable temporal variability and a *multiplicity of time scales*. For instance, *Ohtani et al.* [1992] report from multi-satellite observations that after a period of sluggish growth ($\sim$ 0.5–1.5hr), a near-explosive intensification of the thin cross-tail current sheet occurs during a short interval ($<$ 1 min) just before the onset of the substorm expansion phase. It is also well known that hard X rays emitted during an eruptive flare, which are a reliable signature of prompt electrons, exhibit a sudden transition from a slow growth phase to an impulsive phase. (See, for instance, *Tandberg-Hanssen and Emslie* [1988] and other references therein.) It is clear that any theory of reconnection which proposes to account for the complexity of these phenomena must account for this subtle time-development that includes a sudden transition from a sluggish growth phase to a fast impulsive phase.

Another challenge for theory that has fundamental implications is the inadequacy of the resistive MHD model in the high-S regime when thin current sheets develop. As these thin current sheets become more localized and intense, and the reconnection layer-width Δ falls in the "collisionless" range $d_e \equiv c/\omega_{pe} \ll \Delta \ll d_i \equiv c/\omega_{pi}$ (where ω_{pe} and ω_{pi} are the electron and ion plasma frequencies, respectively), it is not valid to neglect the effect of "collisionless" terms in the generalized Ohm's law

$$\mathbf{E} + \frac{\mathbf{v}\times\mathbf{B}}{c} = \eta\mathbf{J} + \frac{4\pi}{\omega_{pe}^2}\frac{D\mathbf{J}}{Dt} - \frac{\nabla p}{ne} + \frac{\mathbf{J}\times\mathbf{B}}{nec} + \eta_h \nabla^2 \mathbf{J} \ , \quad (1)$$

where $\mathbf{E}$ is the electric field, $\mathbf{B}$ is the magnetic field, $\mathbf{v}$ is the plasma flow velocity, p is the electron pressure (assumed to be a scalar), n is the electron density, e is the magnitude of the electron charge, and $D/dt \equiv \partial/\partial t + \mathbf{v}\cdot\nabla$ is the total convective derivative. The proportionality constant η_h on the right of (1) is commonly referred to as "hyperresistivity" [*Bhattacharjee and Hameiri,* 1986].

In this Letter, we report the results of a numerical simulation of forced (or driven) reconnection with an inward boundary flow, governed by the generalized Ohm's law (1) in a compressible plasma. We contrast our results with those obtained by a resistive MHD simulation (i.e., one in which all terms except the first term on the right of equation (1) are neglected). Two distinctive features emerge in our semi-collisional simulation. First, the maximum thin current sheet amplitude, after a period of slow growth, exhibits a drastic impulsive enhancement at near-Alfvénic growth rates which is insensitive to the value of S. Second, the *geometry* of the reconnection region changes profoundly as the Y-points of the current sheet shrink rapidly to form a structure that approaches an X-point. Consequently, there is a rapid enhancement in the reconnection rate.

Copyright 1996 by the American Geophysical Union.

Paper number 96GL01600
0094-8534/96/96GL-01600$05.00

The rapid acceleration of nonlinear reconnection due to electron pressure gradients was first recognized by *Aydemir* [1992] in the context of kink-tearing instabilities, widely believed to be involved in the so-called "sawtooth collapse" phenomenon in toroidal plasmas. *Wang and Bhattacharjee* [1995] developed an analytical model for this phenomenon, in good accord with Aydemir's simulation results.

We represent the magnetic field as $\mathbf{B} = \hat{\mathbf{z}} \times \nabla\psi(x,y,t) + B_z\hat{\mathbf{z}}$, where $\psi(x,y,t)$ is a flux function. The initial magnetic field is given by $B_x = b_0 \tanh(y/a)$, $B_y = 0$, $B_z = B_0 =$ constant. The compressible MHD equations are

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot (\rho \mathbf{v}) , \tag{2}$$

$$\frac{\partial (\rho \mathbf{v})}{\partial t} = -\nabla \cdot \left[\rho \mathbf{v}\mathbf{v} + \left(p + \frac{B^2}{2}\right)\mathbf{I} - \mathbf{B}\mathbf{B}\right] , \tag{3}$$

$$\frac{\partial p}{\partial t} = -\nabla \cdot (p\mathbf{v}) - (\gamma - 1) p \nabla \cdot \mathbf{v} . \tag{4}$$

Here $\mathbf{I}$ is the unit tensor, $\gamma(= 5/3)$ is the ratio of the specific heats of the plasma, and the variables $\mathbf{x}$, $\mathbf{v}$, t, $\mathbf{B}$, ρ, p, and ψ are nondimensionalized as follows: $\mathbf{B}/b_0 \to \mathbf{B}$, $\mathbf{x}/a \to \mathbf{x}$, $\tau/\tau_A \to t$, $\psi \to \psi/b_0 a$, $v \to v/v_A$, $\rho \to \rho/\rho_0$, $p \to p/(b_0^2/4\pi)$, where ρ_0 is a constant. To complete equations (2)-(4), we need equations for $\psi(x,y,t)$ and $B_z(x,y,t)$. These equations can be obtained by combining Faraday's induction equation with the generalized Ohm's law (1) from which we omit the electron inertia term. The equation for $B_z(x,y,t)$ is

$$\frac{\partial B_z}{\partial t} = -\nabla \cdot (B_z \mathbf{v}) + \mathbf{B} \cdot \nabla v_z + \frac{1}{S} \times \nabla^2 B_z - d_i \left[\nabla \times \left(\frac{\mathbf{J} \times \mathbf{B} - \nabla p}{\rho}\right)\right]_z . \tag{5}$$

If we combine Faraday's equation with the z-component of (1), we obtain

$$\frac{\partial \psi}{\partial t} = -\mathbf{v} \cdot \nabla \psi + \frac{1}{S} J_z + \frac{d_i}{\rho} (\mathbf{J} \times \mathbf{B})_z - \eta_h \nabla^2 J_z . \tag{6a}$$

Equivalently, if we combine Faraday's equation with the component of (1) parallel to $\mathbf{B}$, we obtain

$$\frac{\partial \psi}{\partial t} = -\mathbf{v} \cdot \nabla \psi + \frac{1}{S} \frac{\mathbf{B} \cdot \mathbf{J}}{B_z} - \frac{d_i}{\rho B_z} \mathbf{B} \cdot \nabla p - \frac{\eta_h}{B_z} \nabla^2 J_z , \tag{6b}$$

where $B_z \neq 0$. In writing equations (2)-(6), we have assumed that the ion pressure gradient can be neglected. This assumption can be remedied, and we do not expect it will change our results qualitatively. A more serious assumption is that the electron pressure is taken to be a scalar which is violated in many astrophysical plasmas, but is a common assumption in the Earth's magnetotail and the solar corona. The inclusion of tensor pressure effects in fluid models is complicated by the lack of rigorous closure relations for the off-diagonal elements of the pressure tensor.

In our simulations, we use either equation 6(a) or 6(b), because they are not independent. Indeed, it is a check on the consistency of our code that either of these equations, when combined with equations (2)-(5) yield nearly identical numerical results for the same initial conditions. It should be noted, however, that the "finite-ion-Larmor-radius" terms manifest themselves differently in 6(a) and 6(b). In 6(a), they show up as the Hall term, whereas in 6(b) they show up as the electron pressure term. From (5) and (6b), it is also clear that the time-dependence of B_z, which is neglected in 2D incompressible reconnection dynamics, plays a significant role when collisionless terms are retained in the generalized Ohm's law.

We choose the plasma pressure to satisfy the magnetostatic equilibrium condition, and normalize the pressure so that the plasma beta is equal to one on the boundary $y = \pm 1$. We assume that an electric field, given by $E = E_0(1 + \cos \pi x/2)$, is imposed at the upper and lower boundaries $y = \pm 1$, with $E_0 = 0.01$. This electric field generates an inward boundary flow $v_y (\equiv v_0(1 + \cos kx/2)$, where v_0 is a constant) at the upper and lower boundaries of the simulation domain. We use free-boundary conditions at $x = \pm 2$ for all dependent variables except for the flux function ψ which obeys $d\psi/dt = 0$.

For the initial and boundary conditions discussed above, the dynamical equations preserve the symmetry relation $f(x,y) = \pm f(-x,-y)$ for all times, enabling us to carry out the simulation in half of the physical domain with high spatial resolution. In order to reduce numerical truncation errors and resolve the fine spatial structures, we have developed a Runge-Kutta finite differencing scheme with fourth-order accuracy in both space and time.

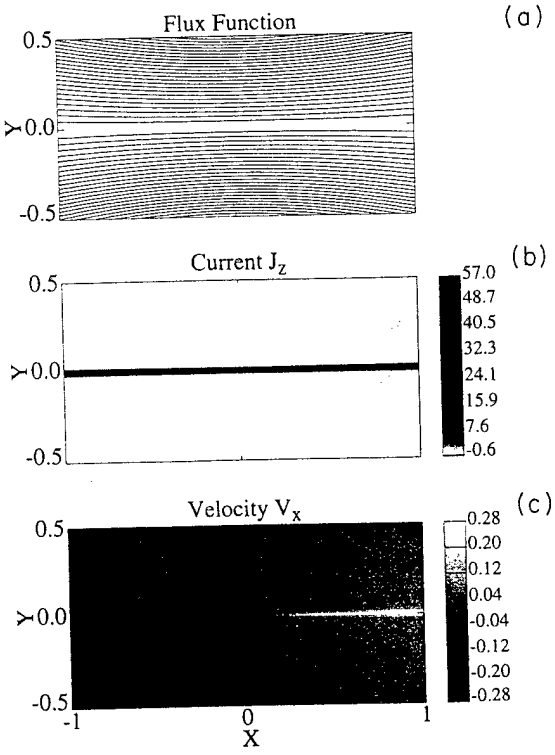

Figure 1. Resistive MHD simulation. Contours of (a) ψ, (b) J_z, and (c) v_x at $t = 20$ for $S = 10^5$.

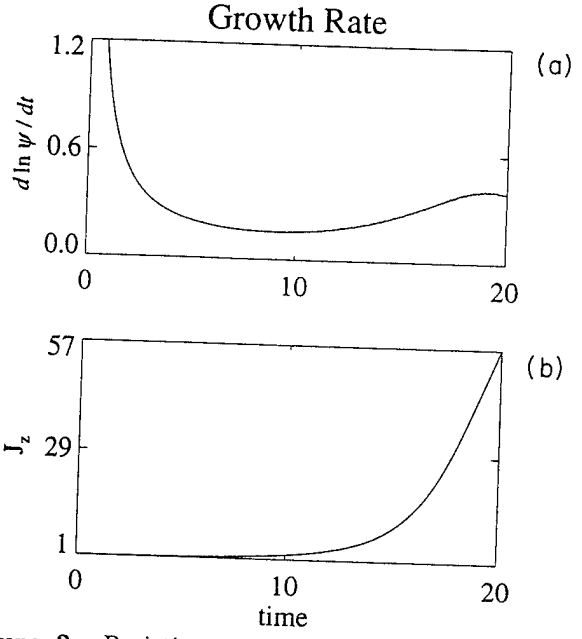

Figure 2. Resistive MHD simulation (with the parameters for Fig. 1). (a) Growth rate of the magnetic flux function ($= d \ln \psi / dt$) at the stagnation point, and (b) the maximum current sheet amplitude as functions of time.

Resistive MHD Simulation

We first present results for an incompressible resistive MHD simulation in which all but the first term on the right of (6a) have been set equal to zero. *Wang et al.* [1996] showed analytically that in this case the peak amplitude of the current sheet (spanning Y-points) exhibits a sudden transition from a sluggish linear growth phase to a rapid nonlinear phase in a characteristic time scale $\tau_N = (\tau_R \tau_0 \tau_A^3)^{1/5}$, where $\tau_0 \equiv a/v_0$ is the time scale of the imposed boundary flow. Figure 1 shows contours of (a) ψ, (b) J_z, and (c) v_x at $t = 20$ for $S = 10^5$. The geometry of this reconnection process conforms to that of Sweet-Parker and *Syrovatsky* [1971]. As shown by *Wang et al.* [1996], the dependence of the reconnection time scale on η is weakened (to $\eta^{-1/5}$ from $\eta^{-1/2}$ in the Sweet-Parker model) due to the flux pileup at the neutral line caused by the inward boundary flows. Figures 2(a) and 2(b) show, respectively, the time-dependence of the growth rate of the magnetic flux function ($= d \ln \psi / dt$) at the stagnation point ($x = 0, y = 0$), and the maximum current sheet amplitude. We note the strong enhancement in the current sheet amplitude, predicted by analysis, as the system passes from the linear to the nonlinear regime. This growth eventually slows down in the nonlinear regime.

Semi-collisional MHD Simulation

With the same initial and boundary conditions, we now present simulation results from the full equations (2)-(6). The electron-pressure term comes into play when the thickness of the current sheet becomes smaller than the ion skin depth. Prior to this time, the magnetic reconnection is controlled by the resistivity and the boundary flow.

Figure 3 shows contours of (a) ψ, (b) J_z, (c) collisionless $E_\parallel (\equiv \nabla_\parallel p/ne)$, and (d) v_x, respectively, for a run with $d_i = 0.1$, $S = 10^5$, and $\eta_h = 5 \times 10^{-10}$ at $t = 20$. Figures 4(a) and 4(b) show, respectively, the time-dependence of the growth rate of the magnetic flux ($= d \ln \psi / dt$) at the stagnation point and the maximum current sheet amplitude. In the linear phase, the current density grows algebraically in time and the growth rate is low, as also seen in the resistive MHD simulation. Driven by inward flows, as the current sheet becomes narrower than the ion skin depth, the pressure gradient (or the Hall term) affects the nonlinear dynamics significantly. Indeed, in the nonlinear stage, the reconnection dynamics is dominated by these collisionless effects. If we compare Figures 2 and 4, we note an impulsive intensification over a few Alfvén times in the growth rate of the magnetic flux as well as the current sheet in Figure 4 (at $t \geq 10$), the rate and magnitude of which is significantly larger than that in Figure 2.

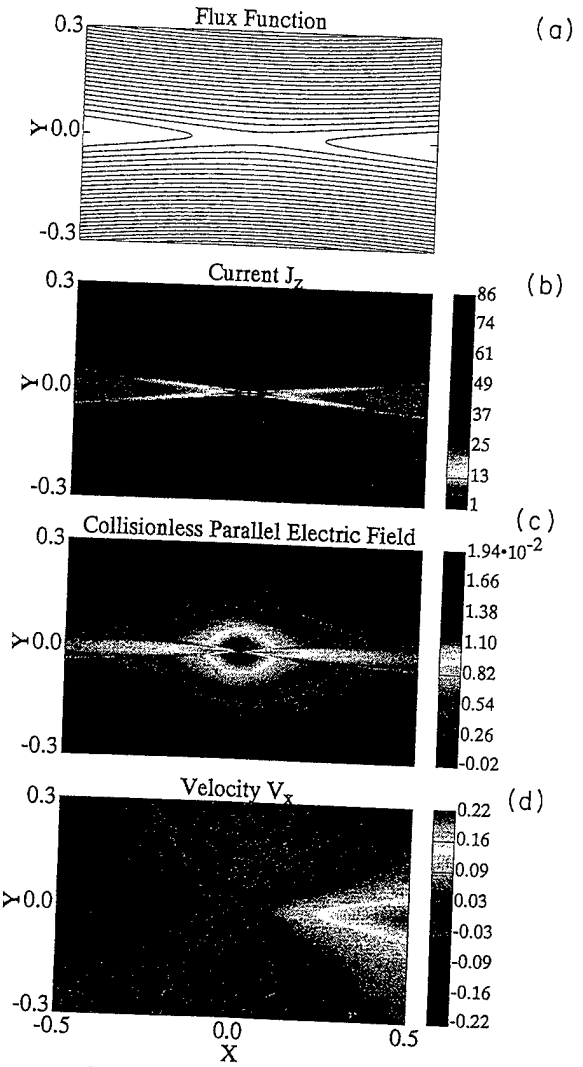

Figure 3. Semi-collisional MHD simulation. Contours of (a) ψ, (b) J_z, (c) collisionless $E_\parallel (\equiv \nabla_\parallel p/ne)$, and (d) v_x at $t = 20$. The initial condition is the same as in Fig. 1, with $d_i = 0.1$, $S = 10^5$, and $\eta_h = 5 \times 10^{-10}$.

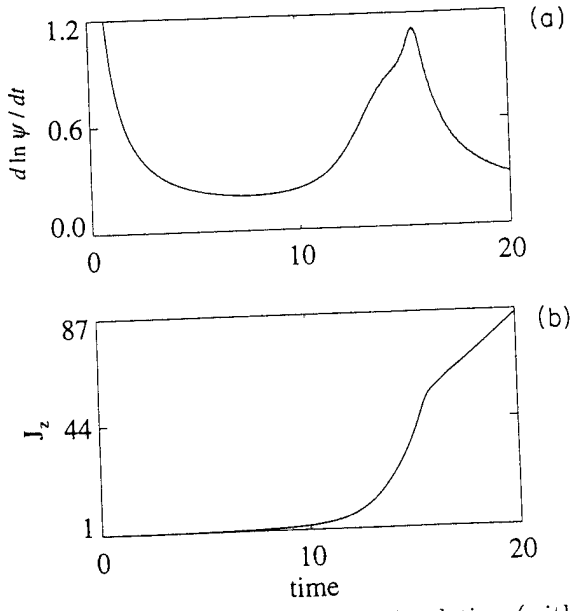

Figure 4. Semi-collisonal MHD simulation (with the parameters for Fig. 3). (a) Growth rate of the magnetic flux function ($= d\ln\psi/dt$) at the stagnation point, and (b) the maximum current sheet amplitude as functions of time.

The growth rate of the magnetic flux and the maximum current sheet amplitude is found to be essentially independent of S (in the range $5 \times 10^4 - 2 \times 10^5$). We expect that the impulsive growth rate is essentially independent of S when S is high, but due to the limitations of computational resources and spatial resolution, we are unable to provide a conclusive S-scaling, valid over a large range of S. In Figure 4(b), we note a transition from impulsive growth to a slower, algebraic growth at $t \sim 16$. This slowing down is primarily due to the effect of hyperresistivity which intervenes when the current sheet becomes very localized. If hyperresistivity is eliminated from the simulation, the impulsive phase persists for longer times.

Electron pressure gradients bring about a profound change in the geometry of the reconnection region. Comparison of Figures 1(b) and 3(b) show that the length of the Y-type current sheet seen in the resistive case is dramatically reduced in the semi-collisional case. The maximum reconnection rate is not only much faster in the semi-collisional case, but also shows a much stronger time-dependence, as it varies from algebraic growth in time to exponential or even near-explosive growth under some conditions [*Wang and Bhattacharjee*, 1995]. Secondly, we emphasize the spatial-scale separation of the collisionless $E_\parallel$ and $J_\parallel$ in Figure 3. Whereas the scale of $E_\parallel$ is of the order of d_i and independent of the resistivity, the scale of $J_\parallel$ is much narrower and controlled by the resistivity. The region with the strongest $E_\parallel$ is located in the vicinity of the stagnation point , but it is exactly zero at the stagnation point . Though the electron pressure term in the generalized Ohm's law does not break field lines, it decouples electron and ion dynamics and prepares the way for the qualitative changes observed in the reconnection dynamics.

The initial geometry used in this simulation is simple, and we do not claim direct applicability to substorms or eruptive flares. As in our recent simulations, we should include the effects of the Earth's dipole field [*Ma et al.*, 1995a] or the geometry of the 2D coronal loop [*Ma et al.*, 1995b] before we attempt to make connections with substorm or coronal phenomenology. Nonetheless, the results presented in this paper are significant because they contain some fundamental features of fast collisionless reconnection that may play a qualitatively important role in the elucidation of such phenomena.

Acknowledgments. We are grateful to Dr. X. Wang for several enlightening discussions. This research is supported by the Air Force Office of Scientific Research grant F49620-93-1-0071 and National Science Foundation grants ATM 93-10157 and ATM 95-29598. Supercomputing resources were provided by the San Diego Supercomputing Center.

References

Aydemir, A. Y., Nonlinear studies of $m = 1$ modes in high-temperature plasmas, *Phys. Fluids B, 4*, 3469, 1992.

Bhattacharjee, A., and E. Hameiri, Self-consistent dynamo-like activity in turbulent plasmas, *Phys. Rev. Lett., 57*, 206, 1986.

Biskamp, D., *Nonlinear Magnetohydrodynamics*, Cambridge University Press, 1993.

Ma, Z. W., C. S. Ng, X. Wang, and A. Bhattacharjee, Dynamics of current sheet formation and reconnection in two-dimensional coronal loops, *Phys. Plasmas, 2*, 3184, 1995a.

Ma, Z. W., X. Wang, and A. Bhattacharjee, Growth, sudden enhancement, and relaxation of current sheets in the magnetotail: Two-dimensional substorm dynamics, *Geophys. Res. Lett., 22*, 2985, 1995b.

Ohtani, S., K. Takahashi, L. J. Zanetti, T. A. Potemra, R. W. McEntire, and T. Iijima, Initial signatures of magnetic field and energetic particle fluxes at tail reconfiguration: Explosive growth phase, *J. Geophys. Res., 97*, 19,311, 1992.

Parker, E. N., *Cosmical Magnetic Fields*, Clarendon Press, Oxford, 1979.

Petschek, H. E., *Magnetic Field Annihilation*, AAS/NASA Symposium on the Physics of Solar Flares, p. 425, edited by W. N. Hess, NASA, Washington, DC, 1964.

Syrovatsky, S. I., Formation of current sheets in a plasma with a frozen-in strong magnetic field, *Sov. Phys. JETP, 33*, 933, 1971.

Tandberg-Hanssen, E., and A. G. Emslie, *The Physics of Solar Flares*, Cambridge University Press, Cambridge, 1988.

Wang, X. and A. Bhattacharjee, Nonlinear dynamics of the $m = 1$ kink-tearing instability in a modified magnetohydrodynamic model, *Phys. Plasmas, 2*, 171, 1995.

Wang, X., Z. W. Ma, and A. Bhattacharjee, Fast reconnection and sudden enhancement of current sheets due to inward boundary flows, to appear in the Special Issue: Invited and Review Papers from the 37th Annual Meeting of the Division of Plasma Physics of the American Physical Society, *Phys. Plasmas*, in press, May, 1996.

A. Bhattacharjee, and Z. W. Ma, Department of Physics and Astronomy, University of Iowa, Iowa City, IA 52242. (e-mail: amitava@iowa.physics.uiowa.edu)

(received December 11, 1995; revised February 29, 1996; accepted March 26, 1996.)

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 101, NO. A2, PAGES 2641–2653, FEBRUARY 1, 1996

A three-dimensional reconnection model of the magnetosphere: Geometry and kinematics

Xiaogang Wang and A. Bhattacharjee

Department of Physics and Astronomy, University of Iowa, Iowa City

Abstract. A fully three-dimensional magnetic reconnection model is proposed for the merging between a spatially nonuniform interplanetary magnetic field and the geomagnetic field. The model generalizes the canonical model of Cowley (1973) and Stern (1973), replacing the two magnetic nulls of the canonical model by a web of A and B nulls and their associated null lines. It is shown that the new A-B model helps resolve some of the difficulties of the canonical model and allows a qualitatively consistent description of magnetic reconnection that begins on the dayside and subsequently evolves to the nightside. A kinematic study based on the induction equation demonstrates that in an ideal plasma, the velocity of the magnetic field has power-law singularities at the separatrices which are the sites of reconnection. Qualitative implications of this global model for ionospheric potential patterns and substorm signatures are discussed.

1. Introduction

The coupling of the solar wind and the interplanetary magnetic field (IMF) to the magnetosphere is one of the most widely studied phenomena in magnetospheric physics. Not only does the solar wind transfer mass, energy, and magnetic flux into the geomagnetic plasma, but a sudden change in the direction of the IMF can cause a major rearrangement of the topology of the magnetosphere, thus causing a substorm "in which a significant amount of energy derived from the solar wind-magnetosphere interaction is deposited in the auroral ionosphere and magnetosphere" [*McPherron*, 1991, p. 600].

Magnetic reconnection is widely believed to play an important role in the phenomena mentioned above. There are many reasons for this belief. Magnetic reconnection can change magnetic topology much faster than diffusion and can transfer mass, energy, and flux between topologically distinct regions. It can also unleash a potentially large source of magnetic free energy, support large induced electric fields that can develop quickly (if circumstances are favorable), and accelerate particles to high energies.

There are two questions that are pertinent to magnetic reconnection: where does it occur and how rapidly? The first is a question about geometry and the second about dynamics. Attempts to answer the second question without knowing the answer to the first can lead to pitfalls. Elsewhere, we have discussed how considerations of geometry as well as dynamics help resolve some of the controversies pertaining to the physics of collisionless tearing instabilities in the magnetotail [*Wang and Bhattacharjee*, 1993; *Bhattacharjee et al.*, 1994; *Harrold et al.*, 1995].

Geometry involves global considerations and leads to the concept that magnetic reconnection (if it is to be distinguished from simple diffusion) occurs in the vicinity of separatrices. For example, an X point or a neutral line in two dimensions or magnetic nulls in three dimensions are sources of separatrices and hence are possible sites of reconnection [*Greene*, 1988; *Lau and Finn*, 1990]. If we know where the separatrices are and how they evolve in time, we know essentially where reconnection occurs as the magnetosphere evolves in time.

There is another point of view which claims that reconnection can occur wherever a parallel electrical field component develops and breaks the ideal magnetohydrodynamic (MHD) constraint. According to this point of view, which leads to the concept of "general magnetic reconnection," separatrices are not required [*Schindler et al.*, 1988]. However, in this paper, we limit our discussion to separatrix-based reconnection since the major IMF-magnetosphere coupling models in the literature, i.e., Dungey's two-dimensional model [*Dungey*, 1961, 1963] or its three-dimensional (axisymmetric) generalization [*Cowley*, 1973; *Stern*, 1973], involve separatrices.

The main goal of this paper is to develop a general geometrical framework for reconnection in the magnetosphere that integrates the dayside with the nightside and to examine what this framework implies qualitatively for ionospheric potential patterns and some local signatures of substorm development, especially during a sudden southward turning of the IMF. The coupling between the IMF and the magnetosphere was first studied in two-dimensional (2D) geometry by *Dungey* [1961, 1963]. In particular, Dungey considered the special case in which the IMF is exactly parallel or antiparallel to

Copyright 1996 by the American Geophysical Union.

Paper number 95JA02184.
0148-0227/96/95JA-02184$05.00

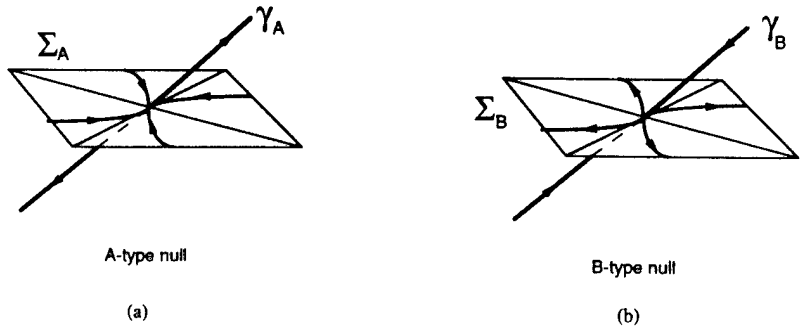

Figure 1. (a) An isolated magnetic null of type A. The Σ_A surface is a stable manifold, and the γ_A line is an unstable manifold. (b) An isolated magnetic null of type B. The Σ_B surface is an unstable manifold, and the γ_B line is a stable manifold.

the Earth's magnetic dipole. *Cowley* [1973] and *Stern* [1973] treated the more general case of the coupling of the Earth's field with an IMF of arbitrary orientation, assuming that the Earth's magnetosphere is axisymmetric and the IMF is uniform. Following *Siscoe* [1988], we shall refer to this model as the canonical model. One of the strengths of the canonical model is that it allows us to consider the geometry of reconnection for a continuum of directions of the IMF, including the cases of exactly northward and southward IMF. However, as discussed by *Siscoe* [1988], the canonical model has a number of difficulties, both theoretical and observational. The present paper attempts to resolve some of these difficulties, and in the process of doing so, we are naturally led to certain qualitative geometrical considerations regarding substorm development.

The following is a plan of the paper. In section 2, we review the canonical model, as it provides a point of departure for further generalizations. In section 3, we present the A-B web model and discuss how it addresses some of the theoretical and observational issues raised by the canonical model. We then discuss the kinematics of magnetostatic reconnection in the context of the A-B web model. Section 4 represents an attempt to identify what might be some of the local signatures for substorms, as the IMF orientation suddenly turns southward from northward. We conclude in section 5 with a discussion of the implications of the new reconnection model and unanswered questions that need further work.

2. The Canonical Model

The canonical model [*Cowley*, 1973; *Stern*, 1973] of magnetic reconnection in the Earth's magnetosphere is embedded in three dimensions and describes magnetic reconnection between the Earth's axisymmetric dipole field and a constant IMF of arbitrary orientation. There are some useful pictures of this process in the paper by *Lau and Finn* [1990] that we reproduce here (in order to keep this paper self-contained).

Magnetic nulls, where the magnetic field vanishes identically, play a crucial role in 3D reconnection phenomena because nulls are the source of separatrices. In what follows, we use the nomenclature of *Cowley* [1973] (and, more recently, of *Greene* [1988] and *Lau and Finn* [1990]) in classifying magnetic nulls. Near a null, the Taylor expansion of $\mathbf{B}$ gives

$$\mathbf{B}(\mathbf{x}) = \nabla\mathbf{B} \cdot \mathbf{x}\,, \tag{1}$$

where $\nabla\mathbf{B}$ is a 3×3 real matrix of zero trace (because $\mathbf{B}$ is a divergenceless field). If $\nabla\mathbf{B}$ is a symmetric matrix (i.e., if the current density $\mathbf{J}$ vanishes at the null), the eigenvalues are real, and the nulls can be either of type A or B. For a type A (B) null, one (two) of the eigenvalues is (are) positive while two are (one is) negative. Figure 1a shows a null of type A. Near the null, the eigenvectors for the negative eigenvalues lie on a two-dimensional plane coinciding with the so-called Σ_A surface which is a stable manifold, whereas the γ_A line that coincides with the eigenvector for the positive eigenvalue is an unstable manifold. If we reverse the arrows in Figure 1a, we obtain a picture of a null of type B, shown in Figure 1b.

Let θ be the angle between the Earth's magnetic dipole and the IMF, with $\theta = \pi$ corresponding to a northward IMF. If the IMF is exactly northward, we obtain the configuration in Figure 2a, with an A-type null above the north pole and a B-type null below the south pole. On the other hand, coupling to an exactly southward IMF ($\theta = 0$) yields an X chain (or a ring of X points) circling from the dayside to the nightside all around the Earth. The projection of this configuration on a meridianal plane yields the picture in Figure 2b, originally envisioned by *Dungey* [1961]. For the case of the northward IMF, the separator that connects the A and B nulls clearly separates the closed geomagnetic lines from the open IMF lines. Forced reconnection (driven by the solar wind) for northward IMF, which we refer to hereafter as reconnection of type N, occurs between the IMF and the closed field lines on the nightside. For the case of the southward IMF, the separator separates the closed geomagnetic field lines from the open IMF lines, but there is a new class of open field lines which have a footpoint in the polar caps. Forced reconnection (driven by the solar wind) for southward

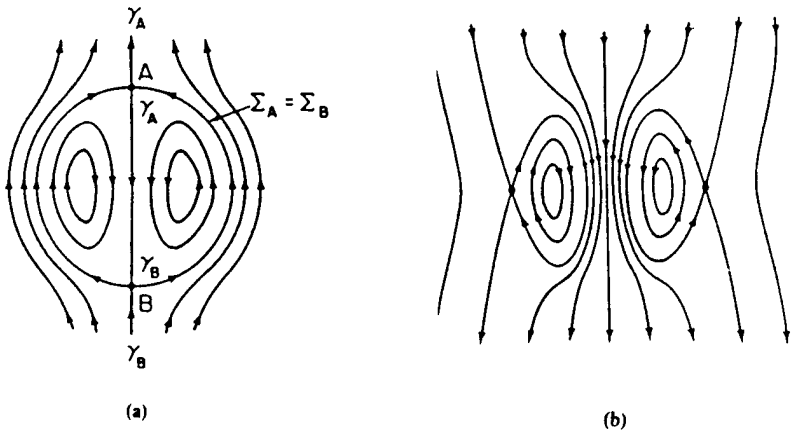

Figure 2. The magnetic geometry when (a) a constant and exactly northward interplanetary magnetic field (IMF) is superposed on an axisymmetric dipole field and (b) a constant and exactly southward IMF is superposed on an axisymmetric dipole field (courtesy of *Lau and Finn* [1990]).

IMF, which we refer to hereafter as reconnection of type S, occurs at first between the IMF and the closed field lines on the dayside, then between the half-open field lines (each with a footpoint on the pole) and the closed field lines on the nightside.

From the discussion above, it can be anticipated that as the IMF turns from northward to southward, the two nulls, initially near the two poles, will move equatorward, depending on the angle between the geomagnetic axis and the direction of the IMF. *Cowley* [1973] and *Stern* [1973] considered the general case of arbitrary IMF orientation. Figures 3a and 3b correspond to the cases $\theta < \pi/2$ and $\theta > \pi/2$, respectively. The Σ_A and Σ_B surfaces that are coincident in Figure 2a become distinct in Figures 3a and 3b. The A-B line, which is referred to as the "single separator" in the canonical model, is the site of magnetic reconnection for an IMF of general orientation, which we refer to, hereafter, as reconnection of type G.

A common feature of all the realizations of the canonical model is the presence of two nulls, one of type A, the other of type B. As discussed by Siscoe in his review [*Siscoe,* 1988, and references therein], the canonical model presents us with a number of theoretical and observational difficulties.

2.1. A Geometric Discontinuity

The first difficulty is that there is a geometric discontinuity in the canonical model. We have noted above that as θ changes from π to zero (IMF changes from northward to southward), the two nulls move equatorward. This process continues till θ reaches a near-zero value. Just at $\theta = 0$, however, these two nulls change to a denumerably infinite number of X points (or a ring of X points) all around the Earth. This sudden and discontinuous change in the geometry is an artifact of the assumption of axisymmetry of the Earth's dipole field.

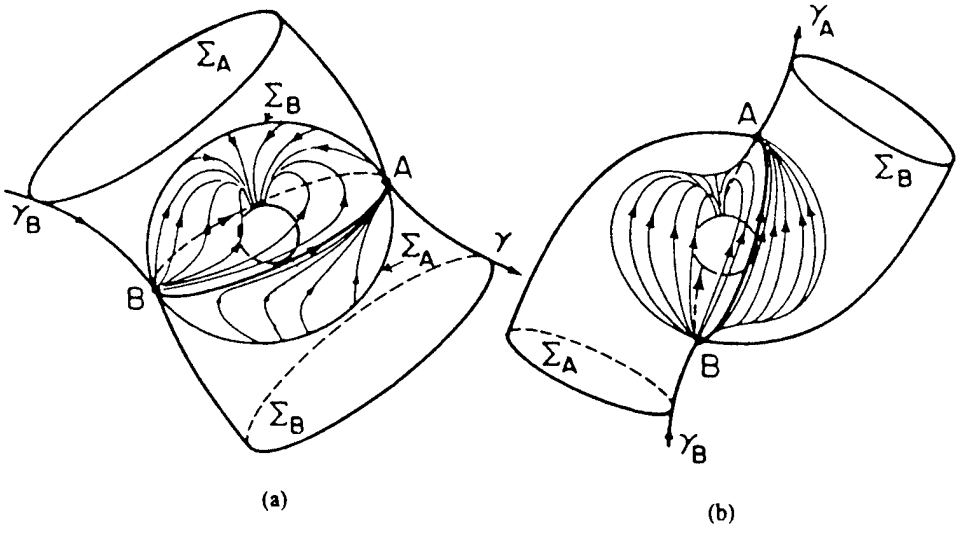

Figure 3. The three-dimensional magnetic geometry when the IMF is oriented (a) at $\theta < \pi/2$ and (b) at $\theta > \pi/2$ (courtesy of *Lau and Finn* [1990]).

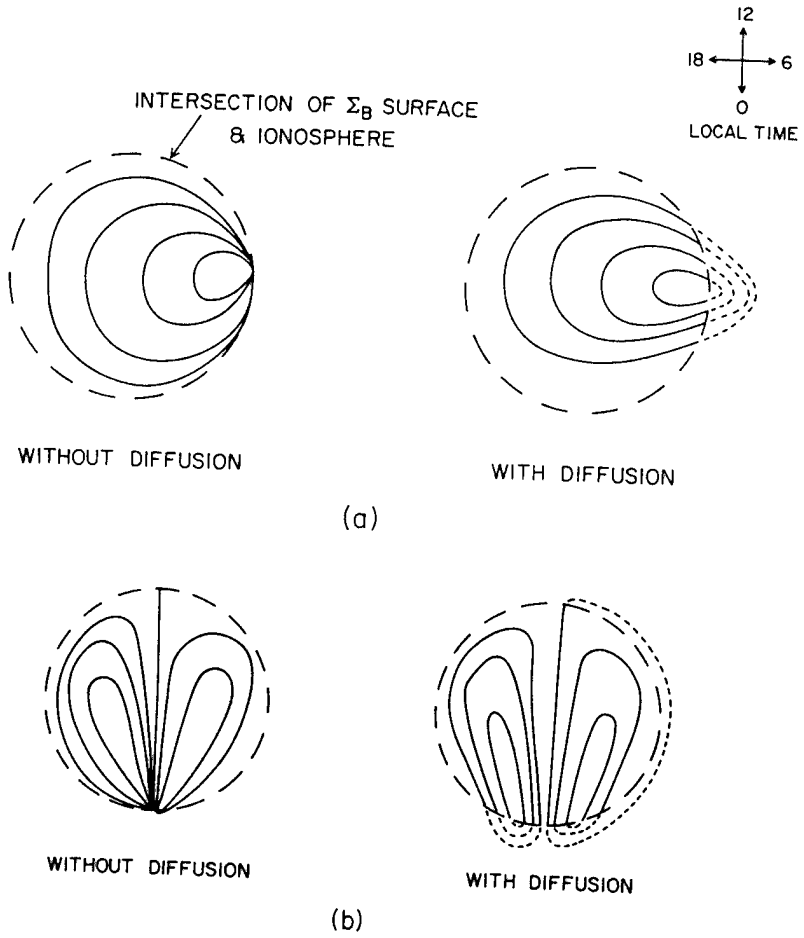

Figure 4. Ionospheric equipotentials for (a) the canonical model with a dayside separator and (b) the canonical model with a dayside A-type null.

2.2. Stern Singularity

The second difficulty, referred to as the Stern singularity by *Siscoe* [1988], is caused by the role of the two nulls as the only topological singularities of the canonical model. For any orientation of the IMF other than exactly southward, all field lines coating the Σ_A (Σ_B) surface intersect at the A null (B null), which, in turn, connects to the ionosphere by a single γ_A (γ_B) line. The solar wind generates a motional electric field in the magnetosphere, governed by the equation $\mathbf{E} + \mathbf{v} \times \mathbf{B}/c = \mathbf{0}$. The equipotentials for this electric field can be mapped along magnetic field lines onto the ionosphere. Then, as pointed out by *Stern* [1973] and explicitly calculated by *Lyons* [1985], the ionospheric equipotentials exhibit a singularity, i.e., a single point where all the equipotentials intersect and the electric field is infinite.

The difficulty is not that this singularity occurs, for it is an intrinsic property of separatrices, but that it prevents the realization of the observed features of the ionospheric auroral equipotential pattern. The Stern singularity is an inevitable consequence of the ideal MHD approximation $\mathbf{E} + \mathbf{v} \times \mathbf{B}/c = \mathbf{0}$. For a real plasma, there must be a small but finite departure from idealness due to diffusion or other collisionless effects such as the electron inertia or anisotropic electron pressure tensor. Not unexpectedly, attempts have been made to resolve the second difficulty (i.e., the Stern singularity) by invoking a diffusion region so that the electric field in the "singular" region becomes finite [*Vasyliunas*, 1984]. A diffusion region does indeed smear out the singularity, but the smeared singularity still casts its shadow in the ionospheric equipotentials which tend to crowd in the vicinity of where the Stern singularity was (see Figure 4a, for example).

2.3. Discordance on the Nightside

As discussed by *Siscoe* [1988, p. 58], whereas "the dayside part" of the diffusion-based models "shows some agreement with observations, the nightside part seems hopelessly discordant." The discordance arises when we consider the expansive phase of the magnetospheric substorm. "In the ionosphere, the explosion of activity signaling the onset of the expansive phase occurs near local midnight and ... the activity remains relatively localized to a few hours of local time around midnight. Nearly unequivocally, the data show that the ionospheric site of dayside merging lies in the polar cap

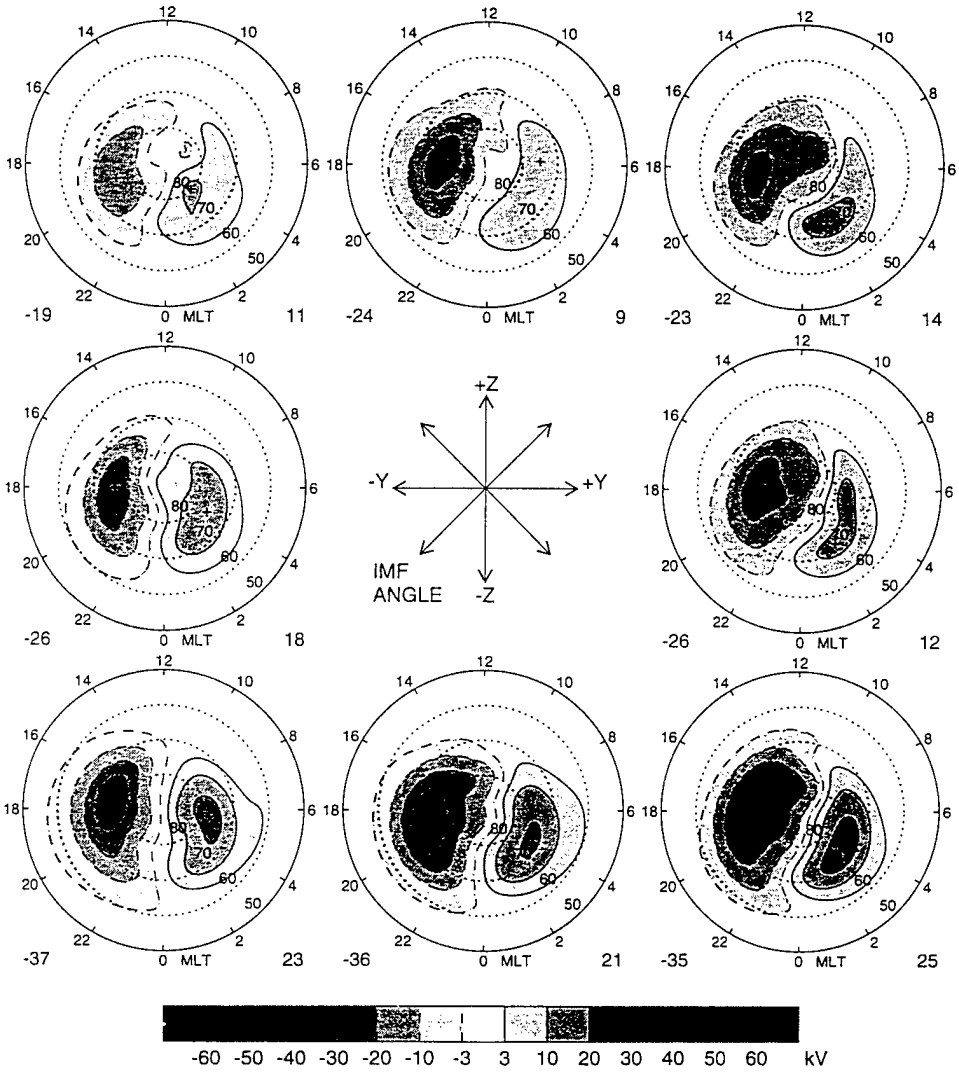

Figure 5. Observed ionospheric equipotential patterns (courtesy of Weimer).

boundary in a region almost diagonally opposite to the point where the substorm expansive phase begins" [*Siscoe*, 1988, p. 58]. However, incorporation of a diffusion layer in the canonical model leaves one with the conclusion that "if nightside merging occurs along the separator line in the canonical separator line model, it must manifest itself in the ionosphere in a region contiguous with the region of dayside merging. This ineluctable property of the model seems in essential conflict with the observations" [*Siscoe*, 1988, p. 58]. In Figures 4a and 4b, we give a qualitative sketch of the equipotential patterns when the projection of the IMF field in the equatorial plane lies in the dawn-dusk and day-night directions, respectively. It is clear from Figures 4a and 4b that the active region rotates with the IMF orientation, which contradicts the observational picture. Typical observed ionosphere electrical equipotential patterns are shown in Figure 5 (reproduced from *Weimer* [1995]) for various IMF orientations. It follows by inspection that the pattern is always broken at local-noon sector regardless of the IMF orientation in the equatorial plane. Hence we are led to the conclusion "One canonical separator line model cannot account for both ... [dayside and nightside] processes" [*Siscoe*, 1988, p. 58].

This difficulty is unavoidable in the canonical model, as shown in Figures 4a and 4b, because the location of the single separator is determined by the IMF orientation. As the IMF orientation with respect to the Earth's field dipole axis changes, the single separator can lie in any local-time sector. Attempts have been made to cure this difficulty by introducing an unbalanced dayside-nightside merging model [*Siscoe and Huang*, 1985]. But as discussed by *Siscoe* [1988], such a model also fails to reproduce the observational feature that the dayside "center-of-action" in the ionosphere is always around the local-noon sector. Alternatively, a "split separator" model, developed by *Crooker* [1985], includes the effect of the Chapman-Ferraro current sheet coating the magnetopause to resolve this difficulty. In this model the A (and the B) null is split into a neutral line terminated by a pair of nulls A′ and A″ (B′ and B″). The A′ and the B′ nulls remain where the original A and B nulls were,

while the A″ and the B″ nulls form a new separator right in the local-noon sector (illustrated by Figure 18 of *Siscoe* [1988]). However, once again one gets into the difficulties (discussed by *Siscoe* [1988]) associated with magnetic merging in such a geometry, because such a geometry is structurally unstable, i.e., the neutral line in 3D can be destroyed by an arbitrarily small perturbation. This issue has been discussed by *Greene* [1988] and *Lau and Finn* [1990] who give a proof that a continuum of nulls, while stable in two dimensions, is structurally unstable in three dimensions.

Before we conclude this section, we comment on other interesting efforts to resolve the difficulties of the canonical model. Notable among them is the "current penetration" model [*Alekseyev and Belen'kaya,* 1983]. This model makes the assumption that the merging line is always perpendicular to the IMF orientation which explains why it is sometimes referred to as a "tangent merging line" model [*Siscoe,* 1988]. However, in a special case such as the coupling to a northward IMF, it is clear by inspection that the location of the merging line is not given correctly. A more recent current sheet model due to *Toffoletto and Hill* [1989, 1993] prescribes a reasonable location for the merging line and predicts qualitatively the observed ionospheric equipotential patterns. It is interesting to note that the Toffoletto-Hill model relies on a special twisted local field configuration to create a dayside neutral merging line and a nightside neutral line. In topological terms, the model is of type S regardless of the IMF orientation.

3. The A-B Web Model

3.1. Geometry

We propose that the canonical model, in which the Earth's magnetic field is axisymmetric and includes two nulls, be generalized to a nonaxisymmetric geometry which includes a "web of nulls and A-B lines" [*Greene,* 1993]. In reality, the real magnetosphere does not have the axisymmetry of the dipole field, nor is the IMF exactly uniform. The axisymmetry of the Earth's dipole field is clearly broken by the depression in the dayside and the stretch in the nightside by the IMF. Observations indicate that spatial nonuniformity is a generic feature of the IMF. (See, for instance, the text by *Parks* [1991] and other references therein.) The spatial nonuniformity can be temporally intermittent, such as that caused by a directional discontinuity, with a typical thickness of about $2\,R_E$, generally observed as it passes a spacecraft in about 30 s at a speed of approximately 420 km/s, about once an hour. The nonuniformity can also have a steady component on a longer spatial scale of 10 R_E or so, caused by the interaction of the solar wind and the magnetosphere. The analytical representation of the magnetic field and the kinematic analysis of the reconnection in the new model is discussed in section 3.2; we begin here with a qualitative description.

The new model is largely motivated by the "split-separator model," which, as pointed out above, is topologically unstable and marks "the end of the road we started on by following the canonical separator line model" [*Siscoe,* 1988, p. 61]. In the new model, the null of A type is broken into an A-B-...-B-A line while the null of B type is broken into an B-A-...-A-B line. Such a configuration is generally structurally stable, with the number of nulls on the A-B-...-B-A and B-A-...-A-B lines determined by the spatial dependence of the magnetic field. The most simple case is realized when the A-null (B-null) of the split separator model is replaced by an A-B-A line (B-A-B line). We repeat for emphasis that whereas in the split-separator model each of the A and B nulls is stretched into a neutral line consisting of infinite numbers of continuous nulls and is terminated by two T-type nulls, in the new model each null is broken into a null line containing a finite number of discrete nulls and terminated by a pair of A (B) nulls, if the original null is A type (B type). Therefore the null line can keep the dayside center of action in the local-noon sector as the split-separator model does but without the difficulties caused by null-splitting: it is structurally stable and has an X point reconnection feature [*Lau and Finn,* 1990; *Greene,* 1993].

For reconnection of type S, the X chain becomes an -A-B-A- null line circling all the way from the dayside to the nightside. Generally, in reconnection of type G (including the type N), the northerly A-B-A line moves toward the south, while the southerly B-A-B line moves toward the north, as shown in Figure 6a. As $\theta \rightarrow 0$, both lines approach the equatorial plane. The A-B-...-B-A line on the duskside makes a half circle from the dayside to the nightside, while the B-A-...-A-B line makes the other half circle (Figure 6b). When $\theta = 0$, the A nulls terminating the A-B-...-B-A line connect with the B nulls terminating the B-A-...-A-B line and thus realize an entire -A-B-A-line on the equatorial plane encircling the globe. Note that we have thus resolved the first difficulty, the geometric discontinuity at the transition from $\theta \rightarrow 0$ to $\theta = 0$, of the canonical model. This resolution becomes possible because the two A and B nulls of the axisymmetric canonical model have now been replaced by a web of A-B nulls.

What happens to the Stern singularity (Figures 4a and 4b)? Since each of the two nulls in the canonical model is transformed into a null line in the new model, one-point singularity is replaced by multiple-point singularities. As discussed in section 2, the presence of such singularities in the ideal limit should not be surprising and is an intrinsic feature of separatrices. In practice, the singularities will be resolved by the presence of a small but finite diffusion (or other nonideal MHD effects) as in the other diffusion models discussed above. In fact, the electric potential drop along a separatrix sustains a reconnection electric field $E_{\|}$. (We elaborate on this point in section 4.) With this $E_{\|}$ field on those separatrix lines, our A-B web model gives an ionospheric equipotential pattern of the type shown in Figure 7. Qualitatively, the picture seems to be in closer accord than Figures 4a and 4b with the observed pattern shown in Figure 5.

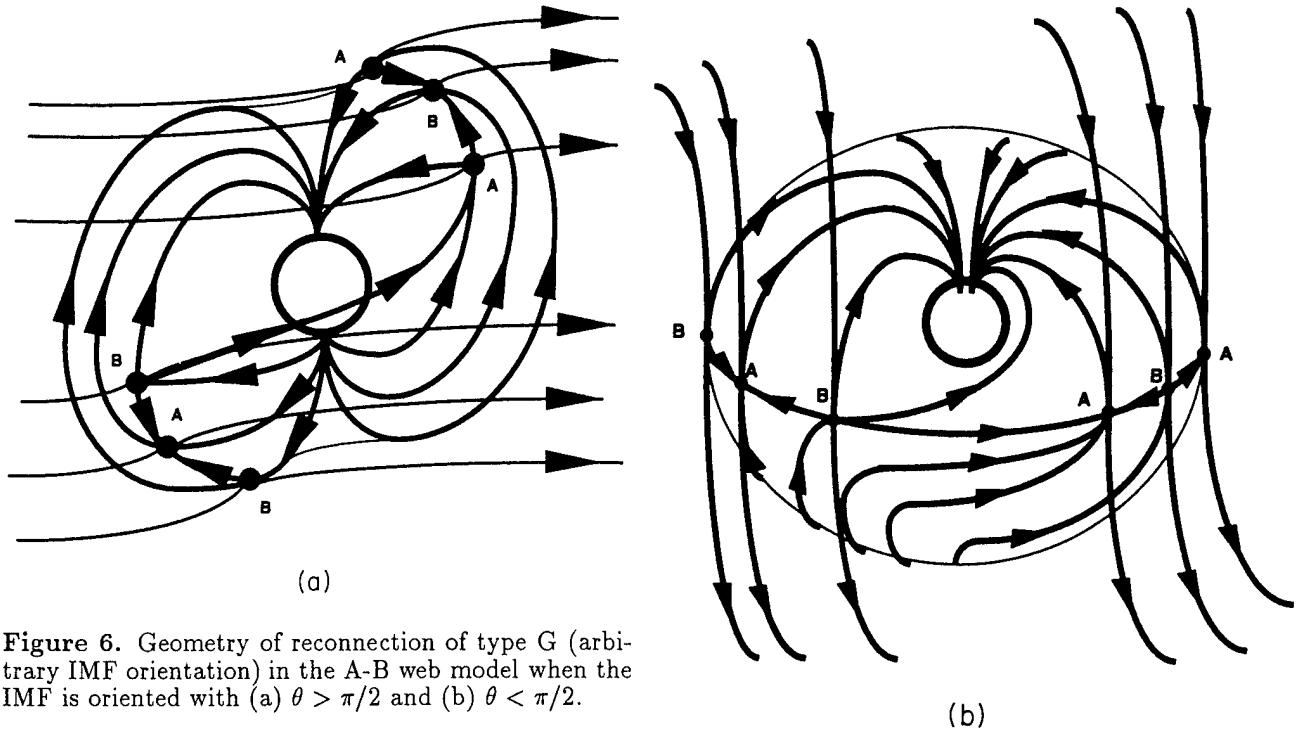

Figure 6. Geometry of reconnection of type G (arbitrary IMF orientation) in the A-B web model when the IMF is oriented with (a) $\theta > \pi/2$ and (b) $\theta < \pi/2$.

Finally, the A-B web model cures, in the same way as the split null-line model (but without the structural instability), the third difficulty of the canonical model: the discordance with the observational fact that most dayside merging activity during the expansive phase of a substorm occurs in the local-noon sector. The difficulty is unavoidable in the canonical model because the location of the single separator is determined by the IMF orientation. As the IMF direction rotates around the Earth's magnetic dipole axis, the separator can be located, in principle, in any local-time sector. However, as shown in Figure 6, on an -A-B-A- web, after the two nulls are split, the "single separator" always moves to the local-noon region just as in the split-separator model. It is on this separator that reconnection is forced by the solar wind. In Figure 7, it is evident that the ionospheric equipotential pattern is broken at the local-noon sector where the A terminator of the separator line is mapped.

3.2. Kinematics of Magnetic Reconnection

For a magnetic field $\mathbf{B}$ embedded in a fluid, $\mathbf{v}$ is called "the velocity of the magnetic field," if it obeys the equation [*Newcomb*, 1958]

$$\frac{\partial \mathbf{B}}{\partial t} - \nabla \times (\mathbf{v} \times \mathbf{B}) = 0 \,. \tag{2}$$

Alternatively, since by Faraday's law,

$$\nabla \times \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t} \,, \tag{3}$$

the restriction imposed by (2) is that the electric field $\mathbf{E}$ can be expressed in the form

$$\mathbf{E} = -\frac{\mathbf{v} \times \mathbf{B}}{c} - \nabla \Phi \,, \tag{4}$$

where Φ is a scalar function. Equation (4) can be solved if we satisfy the solubility condition

$$\mathbf{B} \cdot \nabla \Phi = -\mathbf{E} \cdot \mathbf{B} \,. \tag{5}$$

When a "decent" solution [*Greene*, 1988] of (5) (subject to boundary conditions) exists, the "velocity of the magnetic field" is given by

$$\mathbf{v} = cB^{-2}(\nabla \Phi + \mathbf{E}) \times \mathbf{B} \,. \tag{6}$$

In the context of the present discussion, the sites of magnetic reconnection lie where $\mathbf{v}$ is very large or singular. It is clear that in three dimensions, magnetic nulls are a copious source of separatrices where solutions of (5) generally do not exist.

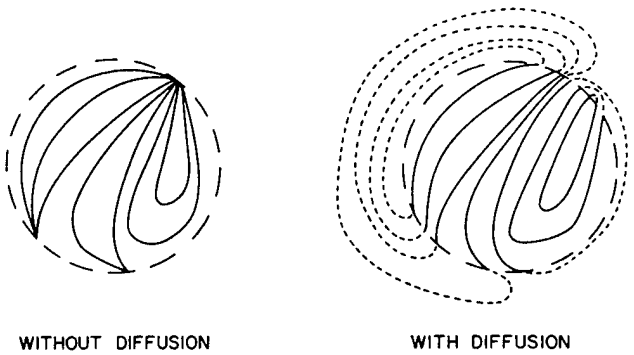

Figure 7. Ionospheric equipotentials for the A-B web model.

These considerations make it clear why the Stern singularity is unavoidable within the framework of the ideal MHD model. Since in ideal MHD, we have $\mathbf{E}+\mathbf{v}\times\mathbf{B}/c=\mathbf{0}$, (5) implies that $\mathbf{B}\cdot\nabla\Phi=\mathbf{0}$. In other words, there is no parallel electric field or a potential drop along a magnetic field line. Since all field lines, except the γ_B line, on the separatrix must pass through the A-type null which is connected to the ionosphere by only one γ_A line, all equipotential lines must then meet at the same singular point at the ionosphere, with the annular region squeezed out of existence. On the other hand, in the A-B web model, all separatrix field lines that lie between neighboring γ_B lines map onto one singular point in the ionosphere (Figure 6a). If the neighboring γ_B lines are generated by a pair of neighboring B-type nulls, the singular point then is the mapping of the A-type null in between the pair of B nulls. However, since field lines between different neighboring nulls map to different singular points, we have multiple singularities in the ionospheric mapping pattern (Figure 7). In practice, the resolution of all such singularities, which are an intrinsic feature of the potential mapping procedure involving a separatrix, is accomplished by noting that there is a nonzero parallel electric field due to nonideal effects in Ohm's law (such as resistivity or finite particle inertia). In the context of the canonical model as well as the A-B web model, this nonideal region smooths out the singularities. (We repeat for emphasis that the problem with the canonical model is not that it needs a diffusion region to resolve the Stern singularity, but rather that even after this is done, one is left with the third difficulty, i.e., the discordance with the observed feature of nightside merging.)

We now consider a specific (local) representation of the magnetic field that allows for the formation of an A-B web. In the standard geocentric solar ecliptic system, $\mathbf{x}$ points from the Earth to the Sun, $\mathbf{y}$ from dawn to dusk, and $\mathbf{z}$ from south to north. Associated with this system is the standard cylindrical coordinate system (R,ϕ,z), with $R=\sqrt{x^2+y^2}$, $\phi=\tan^{-1}(y/x)$, in which we approximate the magnetic field near the merging line $(z=0, R=R_0)$ as

$$B_R = -B_0\frac{z}{R}+\alpha B_0\frac{R-R_0}{R_0}\sin k\phi\,,$$

$$B_\phi = \frac{\alpha}{k}B_0\left(3\frac{R}{R_0}-1\right)\cos k\phi\,,$$

$$B_z = -\lambda^2 B_0\frac{R-R_0}{R_0}+B_0\frac{\alpha z}{R_0}\sin k\phi\,, \qquad (7)$$

where R_0 is a constant and the parameters k, α, and λ are positive and dimensionless, with $\alpha\ll 1$, $\alpha\lambda^2\sim 1$. The model equations (7) can be seen as variant of the TH89 model with selected Fourier components to enable the formation of an A-B line. Recently, there has been some numerical evidence that such a structure appears in three-dimensional MHD simulations of the magnetosphere [*Ogino et al.*, 1994].

Redefining the dimensionless variables

$$\frac{\mathbf{B}}{B_0}\to\mathbf{B}\,,\quad \frac{z}{R_0}\to z\,,\quad \frac{R-R_0}{R_0}\to r$$

and taking $r, z\ll 1$, we have

$$B_r = -z+\alpha r\sin k\phi\,,$$

$$B_\phi = \frac{2\alpha}{k}\cos k\phi\,,$$

$$B_z = -\lambda^2 r+\alpha z\sin k\phi\,, \qquad (8)$$

The magnetic field (8) has the following features: (1) an A-B-A null line at the intersection of the surfaces $r=0$ and $z=0$ (i.e., $R=R_0$ and $z=0$); (2) A-type nulls at $r=0$, $\phi=(2n+1/2)\pi/k$, $z=0$ and B-type nulls at $r=0$, $\phi=(2n-1/2)\pi/k$, $z=0$, where $n=0,\pm1,\pm2,\ldots$; (3) the $z=-\lambda r$ surface as a Σ_A surface and the $z=\lambda r$ as a Σ_B surface. Far away from $r=0$, we obtain the IMF $\mathbf{B}\approx B_z\hat{z}$. For instance, the model can describe qualitatively the geometry of an S-type coupling that includes an -A-B-A- line belt with a width $|z|<1R_E$ at the magnetopause, located roughly at $R=R_0\sim 10R_E$.

The investigation of the dynamics of such a geometry is a complicated problem, beyond the scope of this paper. Instead, we investigate the simpler problem of kinematic reconnection on the separatrices associated with the A-B null lines.

Using (8), the field lines near the A-B null line can be shown to lie at the intersection of a set of level surfaces

$$\left|\frac{z+\lambda r}{z-\lambda r}\right|^{\alpha/\lambda}\tan\left|\frac{k\phi}{2}+\frac{\pi}{4}\right| = C_1\,, \qquad (9)$$

$$\left(z^2-\lambda^2x^2\right)\cos k\phi = C_2 \qquad (10)$$

where C_1 and C_2 are parameters and we have used the approximation $Rd\phi/R_0\cong d\phi$ near the A-B null line.

We now proceed to solve the induction equation (2) for $\mathbf{v}$, given $\mathbf{B}$. Since the electric field is given by $\mathbf{E}=-\nabla\Phi-\partial\mathbf{A}/\partial t$, the condition $\mathbf{E}\cdot\mathbf{B}=\mathbf{0}$ for an ideal plasma gives

$$\mathbf{B}\cdot\nabla\Phi=-\mathbf{B}\cdot\partial\mathbf{A}/\partial t\,. \qquad (11)$$

Following *Lau and Finn* [1990], we add a time-varying vector potential to the time-independent vector potential corresponding to the static $\mathbf{B}$ field. In other words, we write, $\mathbf{A}=\mathbf{A}_0(\mathbf{x})+\mathbf{A}_1(\mathbf{x},t)$ where $\mathbf{B}=\nabla\times\mathbf{A}_0$, and choose

$$\mathbf{A}_1=-\frac{tE_0R_0}{R}\hat{\phi}\,. \qquad (12)$$

Then (4) becomes

$$\mathbf{B}\cdot\nabla\Phi=\frac{B_\phi E_0R_0}{R}\,. \qquad (13)$$

Equation (13) can be integrated, using the method of characteristics, to yield

$$\Phi = \int \frac{dl}{B}\frac{B_\phi E_0 R_0}{R} = \int \frac{Rd\phi}{B_\phi}\frac{B_\phi E_0 R_0}{R}$$

$$= E_0 R_0[\phi - \phi_0(r,\phi,z)]\,, \tag{14}$$

where, choosing $\Phi = 0$ at $r = 0$, we have

$$\phi_0 = \frac{2}{k}\arctan C_1\left[\left|\frac{z+\lambda r}{z-\lambda r}\right|, \phi\right] - \frac{\pi}{2k}\,. \tag{15}$$

Using (6), we calculate the "velocity of the field line" at the separatrices $z = \pm\lambda r$:

$$\mathbf{v} = \frac{2cE_0}{B^2}\left[(\hat{r}B_z - \hat{z}B_r) - \frac{\nabla C_1 \times \mathbf{B}}{k(1+C_1^2)}\right]. \tag{16}$$

Using the relations

$$\frac{\partial C_1}{\partial r} = \frac{2\alpha z C_1}{z^2 - \lambda^2 r^2}\,, \tag{17a}$$

$$\frac{\partial C_1}{\partial z} = -\frac{2\alpha r C_1}{z^2 - \lambda^2 r^2} \tag{17b}$$

we readily see that the velocity has a power-law singularity at the separatrix [*Lau and Finn,* 1990]. This singularity is resolved, in practice, by the presence of a diffusion region near the separatrix where small but finite departures from ideal MHD play a crucial role.

4. Type $N \rightarrow S$ Reconnection and Possible Implications for Substorms

The global model discussed above is qualitative. Yet it is interesting to consider the local implications of this model for substorms. In the absence of a dynamical treatment, the present model cannot provide any information regarding time scales, but it does bring out new possibilities regarding the qualitative evolution of the geometry.

We have considered above the slow and continuous change of the IMF orientation. As this orientation changes smoothly and continuously from northward to southward, the geomagnetic field configuration changes smoothly and continuously from type N, through type G, to type S. However, it is widely believed that most substorm occurrences involve a sudden southward turning of the IMF orientation. During sudden changes of the orientation of the IMF, the topological changes should be sharper and the transitions different than in the continuous case. Consider, for simplicity, the two-dimensional case. If the initial geometry is of type N, with two merging points at the top and the bottom of the magnetosphere, a sudden change of the IMF orientation to the southward direction must lead to the formation of a type-S geometry, with a dayside merging line and a nightside "neutral line." However, the system will pass through stages that mediate the transition between the initial and final states and account for the time required for the sudden change to propagate from the dayside merging line to the nightside neutral line. In this section we will focus our discussion on the transition.

4.1. A Growth Phase Following the Southward Turning of the IMF

Let us assume that initially the IMF is northward. Then there are two distinct null lines at the north and south polar cusp regions. As is well known, the magnetic flux convection is essentially tailward. If the IMF suddenly changes orientation and becomes southward, there must be a neutral-sheet separatrix (NSS) in between the northward and the southward lines of the IMF (Figure 8a). As soon as this NSS reaches the dayside separatrix (the Σ surfaces) of the geomagnetic field, an -A-B-A- null line will form and dayside reconnection, which reduces the dayside flux, will commence. However, on the other side of NSS, the IMF is still northward and the corresponding geometry is still of type N. In this geometry, the merging at the north and south null lines (N&SNLs) is between the IMF field lines and the nightside closed field lines, which also reduces the nightside flux.

The NSS surface will be convected rapidly by the solar wind to the polar cusp region. As shown in Figure 8b, this NSS alters qualitatively the reconnection processes at the north and south null lines. It stops reconnection between the northward IMF and the tail field and starts a new reconnection process on the NSS, transferring flux, particles, and energy into the lobes. This will lead to the completion of a type-S circle: the dayside merging will bring closed dayside flux out of the polar region that will be eventually transferred into the nightside closed field-line region at the N&SNLs. As this flux transfer is completed, there should be an enhancement of the sunward convection, as magnetic flux is lost at the dayside and gained at the nightside in a reconnection process of type S. Thus, as is common knowledge for type-S convection, a velocity shear will tend to build up at both dawn and dusk sides as the process continues. As suggested, for example, by the plasma sheet boundary layer (PSBL) model [*Rostoker and Eastman,* 1987; *Eastman et al.,* 1988], this shear flow will drive the region 1 and the region 2 field-aligned currents. The reconnection process broadly resembles the S-type process discussed earlier; however, it is different from the process that occurs for a steady southward IMF in that it provides a possible mechanism for particle, flux, and energy input from the solar wind into the tail in the near-Earth sector of the tail where the N&SNL merging lines are located. We classify this type of reconnection as north-to-south ($N \rightarrow S$) type. An important consequence of this N-S type reconnection is that the IMF flux and solar wind particles will be brought into the central plasma sheet from both the dawn and dusk sides of the near-Earth plasma sheet, since reconnection occurs dominantly on the separator connecting the north A-null terminator and the south B-null terminator of the null lines. This may be a possible explanation for the observed solar wind par-

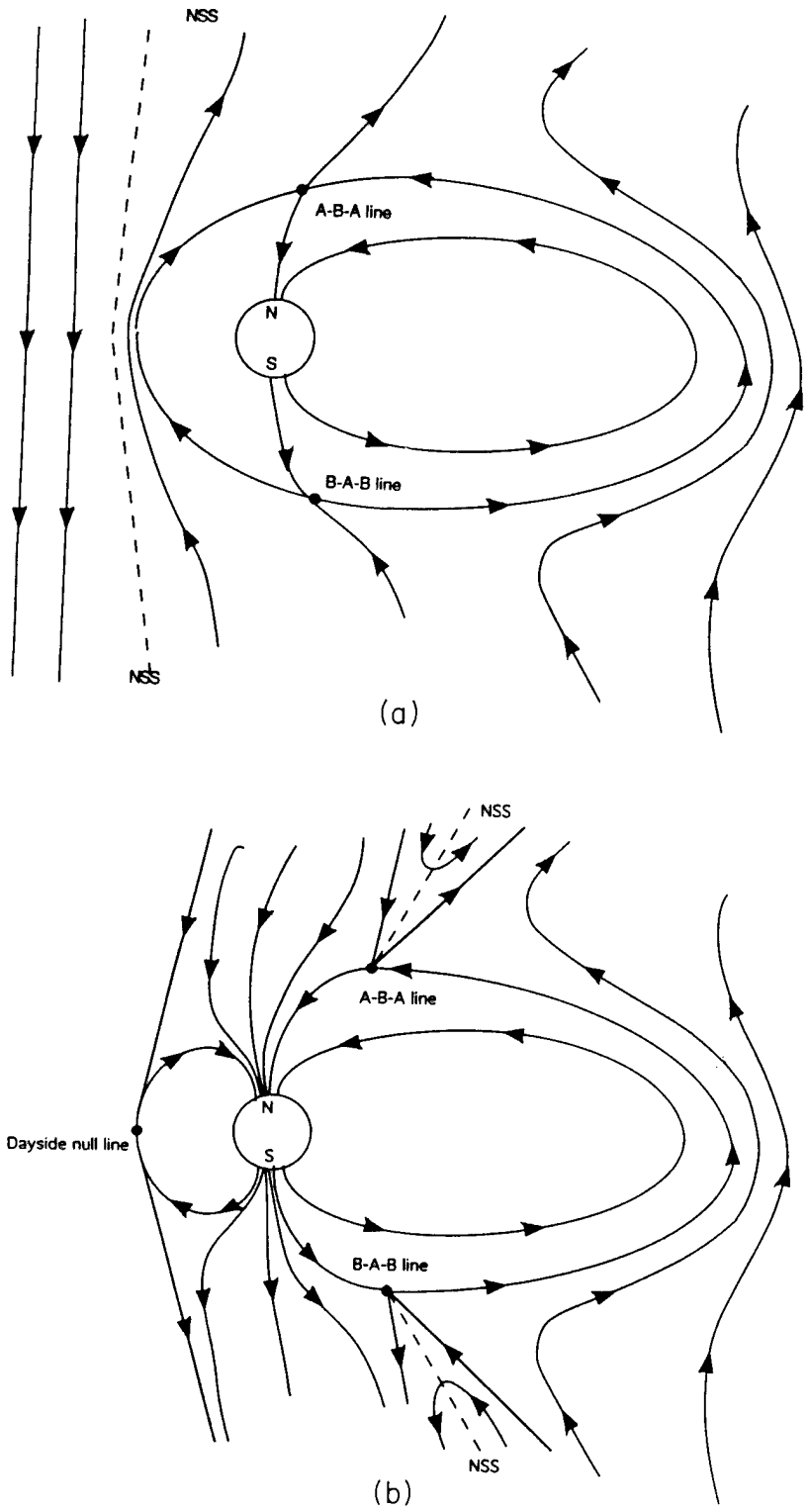

Figure 8. Evolution of the neutral-sheet separatrix (NSS) as the IMF turns from northward to southward. (a) The beginning of the process at the dayside, (b) a second stage of the process when the NSS has moved beyond the polar cusp region, and (c) a later stage as the current sheet thins are shown.

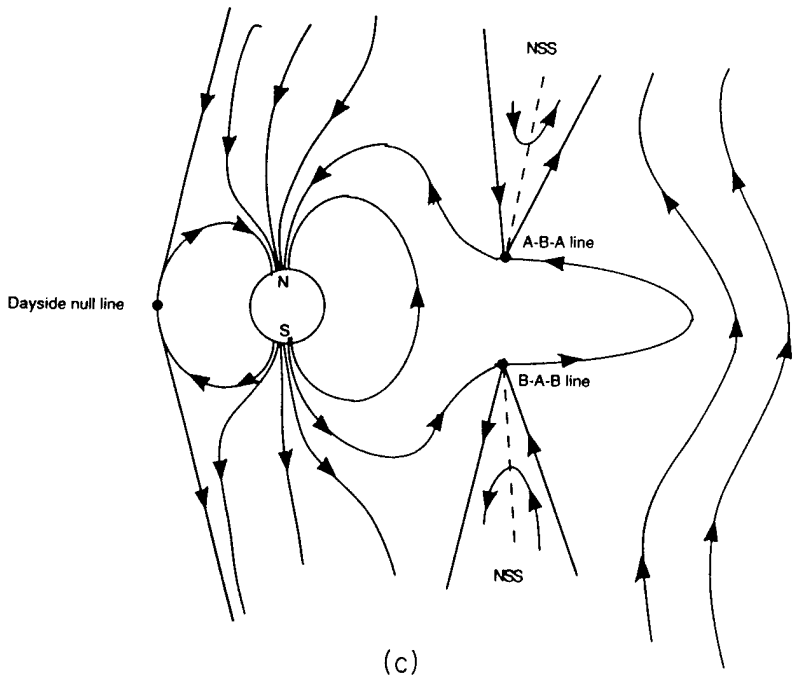

Figure 8. (continued)

ticle component in the central plasma sheet during the growth phase [*Lennartsson,* 1992; *Huang et al.,* 1992].

4.2. The Onset of the Expansion Phase

As the NSS moves further downstream, the geometry has to eventually evolve into a configuration of type S. During this process, the cross-tail current sheet keeps thinning further, driven by the solar wind (Figure 8c). For magnetic reconnection driven by inward boundary flows, the reconnection rate is mostly controlled by the driving electric field and rather weakly by the dissipation mechanism [*Sato and Hayashi,* 1979; Wang et al., Fast reconnection and sudden enhancement of current sheets due to inward boundary flows, submitted for a Special Issue of *Phys. Plasmas*: Invited and Review Papers from the 37th Annual Meeting of the Division of Plasma Physics of the American Physical Society, 6-10 November 1995, 1995; hereinafter referred to as Wang et al., submitted manuscript, 1995].

It is widely believed that the current sheet thinning leads to the onset of the expansion phase at near-Earth distances, though there is, as yet, no definitive identification of the precise physical mechanism for onset. Current-driven instabilities that can cause current disruption and diversion [*Chao et al.,* 1977; *Lui,* 1979; *Lui et al.,* 1988, 1992], ballooning instabilities driven by plasma pressure gradients in regions of unfavorable curvature [*Roux et al.,* 1991; *Erickson and Heinemann,* 1994], and magnetosphere-ionosphere coupling [*Kan et al.,* 1988; *Kan,* 1993] are among the mechanisms that are currently being considered.

In the context of the near-Earth neutral line (NENL) model [*Hones,* 1979], it has been suggested repeatedly that collisionless tearing instabilities can provide a possible mechanism for onset. This subject has a long and controversial history, and it has gradually become clear from independent analytical work and electromagnetic particle simulations by several groups that collisionless tearing modes are strongly stabilized by the presence of a B_z field. (See, for instance, the recent papers by *Pellat et al.* [1991], *Wang and Bhattacharjee* [1993], *Bhattacharjee et al.* [1994], and other references therein.) Nonetheless, it is possible that in some circumstances the B_z field may be reduced by strong externally driven flows from the boundary or if the equilibrium B_y field has a strong spatial dependence [*Wang and Bhattacharjee,* 1993]. Under these conditions, it may be possible to see a sudden transition to a fast growth of thin current sheets (Wang et al., submitted manuscript, 1995). As to how the magnetotail with an embedded thin current sheet disrupts remains an open question. But assuming that one of the mechanisms mentioned above is operative, we proceed with our present discussion after the substorm onset and dipolarization have already occurred.

4.3. The Late Expansion/Recovery Phase

The two N&SNLs will move equatorward further and eventually form a nightside null line just before or in this late phase. The nightside null line may be expected to eventually cut the closed tail field to form a shorter tail and release a plasmoid, which is one of the significant predictions of the NENL model. However, in the presence of a significant B_y field (see, for example, *Akasofu et al.* [1978]), the precise topological characteristics of this plasmoid are likely to be much more complex than originally envisioned. Perhaps it is more appropriate to regard such structures as flux ropes with significant field-aligned currents [*Elphic et al.,* 1986; *Paranicas and Bhattacharjee,* 1989]. The newly formed null line will

move downstream further, and the configuration will tend to the one associated with S-type coupling.

5. Discussion

The main goal of this paper is the development of a three-dimensional geometrical framework for understanding dayside as well as nightside reconnection. As discussed in sections 2 and 3, this framework is a natural generalization of the canonical model and its variants which have provided much of the basis for our present understanding of observations as well as three-dimensional MHD simulations. We are motivated to seek this generalization not merely to pursue generality for its own sake but because some important observational features are not adequately addressed by the earlier models. In particular, we have considered qualitatively, but in some depth, the observations pertaining to the equipotential patterns in the ionosphere.

The canonical model with two nulls is generalized to an A-B web model with multiple nulls and separatrices. We have shown that this generalization resolves some of the outstanding difficulties of the canonical model, one of which has to do with the signatures of nightside reconnection in the expansion phase of a substorm. It has been our guiding principle that the construction of a global geometrical framework must include at once both the dayside and the nightside geometry and that it would be difficult to develop a model for substorms without such an effort. In particular, since magnetic nulls are the source of the A-B web, it is necessary to examine magnetic field data from observations as well as computer simulations to identify magnetic nulls. It is interesting to note that there is clear evidence of null lines in both northern and southern hemispheres in the recent studies of *Kaymaz et al.* [1994a, 1994b] based on IMP 8 data.

Our model is merely kinematic and does not provide any information on self-consistent plasma dynamics. We see it as a first step toward defining the correct geometry on which further dynamical studies should be based. A dynamical study, which is likely to involve extensive computation, should provide quantitative information on how reconnection at N&SNLs can transfer particles, energy, and flux into the tail region and how rapidly the current sheet thins. This may also pave the way for a clearer understanding of the possible mechanisms that can trigger the onset. We hope to explore these issues in future work, based on numerical simulations.

Acknowledgments. We are grateful to C. Y. Huang, G. L. Siscoe, and V. Vasyliunas for helpful discussions. This work is supported by the Air Force Office of Scientific Research grant F49620-93-1-0071 and the National Science Foundation grant ATM-9209006.

The Editor thanks J. Birn and another referee for their assistance in evaluating this paper.

References

Akasofu, S.-I., A. T. Y. Lui, C. I. Meng, and M. Haurwitz, Need for a three-dimensional analysis of magnetic fields in the magnetotail during substorms, *Geophys. Res. Lett., 5,* 283, 1978.

Alekseyev, I. I., and Y. S. Belen'kaya, Electric field in an open model of the magnetosphere, *Geomagn. Aeron.,* Engl. Transl., *23,* 57, 1983.

Bhattacharjee, A., X. Wang, and M. Tanaka, Collisionless tearing instabilities and their possible implications for substorms, in *Substorms 2, Proceedings of the Second International Conference on Substorms*, p. 121, Univ. of Alaska, Fairbanks, 1994.

Chao, J. K., J. R. Kan, A. T. Y. Lui, and S.-I. Akasofu, A model for thinning of plasma sheet, *Planet. Space Sci., 25,* 703, 1977.

Cowley, S. W. H., A qualitative study of the reconnection between the Earth's field and interplanetary field of arbitrary orientation, *Radio Sci., 8,* 903, 1973.

Crooker, N. U., A split separator line merging model of the dayside magnetopause, *J. Geophys. Res., 90,* 12,104, 1985.

Dungey, J. W., Interplanetary magnetic field and the auroral zones, *Phys. Rev. Lett., 6,* 47, 1961.

Dungey, J. W., The structure of the exosphere or adventures in velocity space, in *Geophysics. The Earth's Environment,* Gordon and Breach, New York, 1963.

Eastman, T. E., G. Rostoker, L. A. Frank, C. Y. Huang, and D. G. Mitchell, Boundary layer dynamics in the description of magnetospheric substorms, *J. Geophys. Res., 93,* 14,411, 1988.

Elphic, R. C., C. A. Cahill, K. Takahashi, S. J. Bame, and C. T. Russell, ISEE-1 and ISEE-2 observations of magnetic flux ropes in the magnetotail: FTEs in the plasma sheet?, *Geophys. Res. Lett., 13,* 648, 1986.

Erickson, G. M., and M. Heinemann, A mechanism for magnetospheric substorms, in *Substorms 2, Proceedings of the Second International Conference on Substorms,* p. 333, Univ. of Alaska, Fairbanks, 1994.

Greene, J. M., Geometrical properties of three-dimensional reconnecting magnetic fields with nulls, *J. Geophys. Res., 93,* 8583, 1988.

Greene, J. M., Reconnection of vorticity lines and magnetic lines, *Phys. Fluids B, 5,* 2355, 1993.

Harrold, B. G., A. Bhattacharjee, and X. Wang, Tearing stability of the two-dimensional magnetotail, *Phys. Plasmas,* in press, 1995.

Hones, E. W., Jr., Transient phenomena in the magnetotail and their relation to substorms, *Space Sci. Rev., 23,* 393, 1979.

Huang, C. Y., L. A. Frank, G. Rostoker, J. Fennell, and D. G. Mitchell, Nonadiabatic heating of the central plasma sheet at substorm onset, *J. Geophys. Res., 97,* 1481, 1992.

Kan, J. R., A global magnetosphere-ionosphere coupling model of substorms, *J. Geophys. Res., 98,* 17,263, 1993.

Kan, J. R., L. Zhu, and S.-I. Akasofu, A theory of substorms: Onset and subsidence, *J. Geophys. Res., 93,* 5624, 1988.

Kaymaz, Z., G. L. Siscoe, N. A. Tsyganenko, and R. P. Lepping, Magnetotail views at 33 R_E: IMP 8 magnetometer observations, *J. Geophys. Res., 99,* 8705, 1994a.

Kaymaz, Z., G. L. Siscoe, J. G. Luhmann, R. P. Lepping, and C. T. Russell, Interplanetary magnetic field control of magnetotail magnetic field geometry: IMP 8 observations, *J. Geophys. Res., 99,* 11,113, 1994b.

Lau, Y.-T., and J. M. Finn, Three-dimensional kinematic reconnection in the presence of field nulls and closed field lines, *Astrophys. J., 350,* 672, 1990.

Lennartsson, W., A scenario for solar wind penetration of Earth's magnetic tail based on ion composition data from the ISEE 1 spacecraft, *J. Geophys. Res., 97,* 19,221, 1992.

Lui, A. T. Y., Observations of plasma sheet dynamics during magnetospheric substorms, in *Dynamics of the Magnetosphere,* edited by S.-I. Akasofu, p. 563, D. Reidel, Norwell, Mass., 1979.

Lui, A. T. Y., R. E. Lopez, S. M. Krimigis, R. W. McEntire, L. J. Zanetti, and T. A. Potemra, A case study of magnetotail current disruption and diversion, *Geophys. Res. Lett., 15,* 721, 1988.

Lyons, L. R., A simple model for polar cap convection patterns and generation of q auroras, *J. Geophys. Res., 90,* 1561, 1985.

McPherron, R. L., Physical processes producing magnetospheric substorms and magnetic storms, *Geomagnetism, 4,* 593, 1991.

Newcomb, W. A., Motion of magnetic lines of force, *Ann. Phys. N.Y., 3,* 347, 1958.

Ogino, T., R. J. Walker, and M. Ashour-Abdalla, The effects of IMF orientation on the polar cap and magnetosphere, *Eos Trans. AGU, 75*(44), Fall Meet. Suppl., 536, 1994.

Paranicas, C., and A. Bhattacharjee, Relaxation of magnetotail plasmas with field-aligned currents, *J. Geophys. Res., 94,* 479, 1989.

Parks, G. K., *Physics of Space Plasmas,* Addison-Wesley, Reading, Mass., 1991.

Pellat, R. F., F. V. Coroniti, and P. L. Pritchett, Does ion tearing exist?, *Geophys. Res. Lett., 18,* 143, 1991.

Rostoker, G., and T. E. Eastman, A boundary layer model for magnetospheric substorms, *J. Geophys. Res., 92,* 12,187, 1987.

Roux, A., S. Perraut, A. Morane, P. Robert, A. Korth, G. Kremser, A. Pederson, and Z. Y. Pu, Role of the near-Earth plasma sheet at substorms, in *Magnetospheric Substorms, Geophys. Monogr. Ser.,* Vol. 64, edited by J. R. Kan, T. A. Potemra, S. Kokubun, and T. Iijima, p. 201, AGU, Washington, D.C., 1991.

Sato, T, and T. Hayashi, Externally driven magnetic reconnection and a powerful magnetic energy converter, *Phys. Fluids, 22,* 1189, 1979.

Schindler, K., M. Hesse, and J. Birn, General magnetic reconnection and helicity, *J. Geophys. Res., 93,* 5547, 1988.

Siscoe, G. L., The magnetospheric boundary, in *Physics of Space Plasmas (1987),* edited by T. Chang, G. B. Crew, and J. R. Jasperse, Sci. Publ., Cambridge, Mass., 1988.

Siscoe, G. L., and T. S. Huang, Polar cap inflation and deflation, *J. Geophys. Res., 90,* 543, 1985.

Stern, D. P., A study of the electric field in an open magnetospheric model, *J. Geophys. Res., 78,* 7292, 1973.

Toffoletto, F. R., and T. W. Hill, Mapping of solar wind electric field to the Earth's polar caps, *J. Geophys. Res., 94,* 329, 1989.

Toffoletto, F. R., and T. W. Hill, A non-singular model of open magnetosphere, *J. Geophys. Res., 98,* 1339, 1993.

Vasyliunas, V. M., Steady state aspects of magnetic field line merging, in *Magnetic Reconnection,* in *Space and Laboratory Plasmas,* Geophys. Monogr. Ser., Vol. 30, edited by E. W. Hones Jr., p. 25, AGU, Washington, D.C., 1984.

Wang, X., and A. Bhattacharjee, Global asymptotic equilibria and collisionless tearing stability of magnetotail plasmas, *J. Geophys. Res., 98,* 19,419, 1993.

Weimer, D., Models of high-latitude electric potentials derived with a least-error fit of spherical harmonic coefficients, *J. Geophys. Res., 100,* 19,595, 1995.

A. Bhattacharjee and X. Wang, Department of Physics and Astronomy, University of Iowa, Iowa City, IA 52242. (e-mail: amitava@iowa.physics.uiowa.edu; xwang@iowa.physics.uiowa.edu)

(Received November 21, 1994; revised July 10, 1995; accepted July 10, 1995.)

Nonlinear dynamics of the $m=1$ kink-tearing instability in a modified magnetohydrodynamic model

Xiaogang Wang and A. Bhattacharjee
Department of Physics and Astronomy, University of Iowa, Iowa City, Iowa 52242

(Received 19 May 1994; accepted 23 September 1994)

A theory is given for the nonlinear dynamical evolution of the collisionless $m=1$ kink-tearing instability, including the effects of electron inertia and electron pressure gradient in a generalized Ohm's law. It is demonstrated that electron pressure gradients can cause near-explosive growth in the nonlinear regime of a thin $m=1$ island. This near-explosive phase is followed by a rapid decay phase as the island width becomes comparable to the radius of the sawtooth region. An island equation is derived for the entire nonlinear evolution of the instability, extending recent work on the subject [X. Wang and A. Bhattacharjee, Phys. Rev. Lett. **70**, 1627 (1993)] to include the effects of both electron inertia and electron pressure gradient. Comparisons are made with experimental data from present-day tokamaks. It is suggested that the present model not only accounts for fast sawtooth crashes, but also provides possible explanations for the problems of sudden onset and incomplete reconnection that have been, heretofore, unexplained features of observations. © *1995 American Institute of Physics.*

I. INTRODUCTION

The $m=1$ kink-tearing instability in tokamaks is distinct from the $m>1$ tearing instabilities. In resistive magnetohydrodynamics (MHD), Rutherford[1] showed that the nonlinear $m>1$ modes evolve in time according to the relation $w\sim\eta t$, where w is the size of the island, η is the resistivity, and t is time. Nonlinear $m=1$ modes behave differently and are characterized by a current sheet at the separatrix. The current sheet was seen in several numerical simulations,[2–7] and has been described analytically by Waelbroeck.[8] If the current sheet is neglected, one might infer that the nonlinear mode grows exponentially.[9] However, Waelbroeck has demonstrated that the presence of the current sheet in a Y-shaped separatrix slows down the nonlinear growth of the mode to an algebraic rate, given by the relation $w\sim\eta t^2$, consistent with the predictions of the Sweet–Parker[10] and Kadomtsev[11] models. Recently, Biskamp[12] has extended Waelbroeck's analytical results, and shown that the theory is consistent with high-Lundquist-number numerical simulations.

Free reconnection due to the $m=1$ instability is very similar to forced reconnection, in that both processes are "helicity conserving."[8,13] Forced reconnection, which is widely believed to occur in laboratory and space plasmas, can also produce current sheets.[4,5,8,13] This makes the theoretical study of the $m=1$ instability all the more interesting, because it represents a useful paradigm for fast reconnection in high-temperature plasmas.

Although the nonlinear dynamics of the $m=1$ instability is well understood in the resistive MHD model, the predictions of the theory do not account for significant features of sawtooth oscillations in tokamaks. There are three principal observational features of sawtooth oscillations that pose a challenge for theory: the first is the *rapidity* of the sawtooth collapse; the second is the *sudden onset* of the collapse; and the third is the *incomplete relaxation of the current profile* during which the central q value remains nearly fixed at a value lower than unity while the temperature profile relaxes completely.

In the 1970s, essential features of the available data were thought to be explained by the Kadomtsev model.[11] Kadomtsev showed that in the nonlinear phase of the $m=1$ kink-tearing mode, reconnection occurs at the separatrix on the characteristic Sweet–Parker time scale[10] $\tau_K=(\tau_A\tau_R)^{1/2}$, where τ_A is the Alfvén time scale and τ_R is the resistive diffusion time scale. For most tokamaks operating during the 1970s, the time scales τ_A and τ_R were typically of the order of 10^{-7} and 10^{-1} s, respectively, which gave $\tau_K\sim 100$ μs, in agreement with the sawtooth collapse time then observed. In larger and hotter tokamaks, such as the Joint European Tokamak (hereafter JET)[14] and the Thermonuclear Fusion Text Reactor (hereafter TFTR),[15] the Kadomtsev model predicts a time scale which is one to two orders of magnitude larger than the observed collapse time. (It is not uncommon in these devices to obtain collapse times in the range 20–100 μs, whereas $\tau_K\sim 2-10$ ms.) It is possible to bring the theoretical prediction of the collapse time closer to observations by invoking a resistivity larger than classical in the reconnection layer, but since the resistive $m=1$ kink-tearing mode slows down nonlinearly and exhibits algebraic growth with time, it would still not be possible to account for the onset or "fast trigger" of the sawtooth collapse.[16] These discrepancies between the theoretical predictions of resistive MHD and observations have stimulated considerable theoretical research. While it is widely believed that the nonlinear $m=1$ instability in the collisionless regime is a strong candidate for the observed features of sawteeth, there is as yet no consensus on what should be the appropriate physical ingredients of a correct theoretical model of the $m=1$ instability. In other words, if we write the generalized Ohm's law in the form

$$\mathbf{E}+\mathbf{u}\times\mathbf{B}/c=\mathbf{R} \tag{1}$$

(where $\mathbf{E}$ is the electric field, $\mathbf{B}$ is the magnetic field, $\mathbf{u}$ is the plasma velocity, and c is the speed of light), opinions appear

 1070-664X/95/2(1)/171/11/$6.00 © 1995 American Institute of Physics

to be divided on what physical effects **R** must include in order to account for the salient features of observations.

Wesson[17] has proposed a heuristic modification of the Kadomtsev scaling,[11] based on the form $\mathbf{R}=\eta\mathbf{J}+(4\pi/\omega_{pe}^2)\mathbf{u}\cdot\nabla\mathbf{J}$. Wesson's scaling predicts much faster reconnection rates than Kadomtsev's, and is apparently in good agreement with experimental results. In contrast, Drake and Kleva[18] have shown by numerical simulation that if $\mathbf{R}=\eta\mathbf{J}+(4\pi/\omega_{pe}^2)D\mathbf{J}/Dt$ (where $D/Dt=\partial/\partial t+\mathbf{u}\cdot\nabla$), then in the limit $\eta\rightarrow 0$, current sheets tend to form on spatial scales smaller than the electron collisionless skin depth $d_e=c/\omega_{pe}$ (where c is the speed of light and ω_{pe} is the plasma frequency), which slows down the reconnection rate. Thus, Drake and Kleva infer that electron inertia, *by itself*, cannot produce a fast sawtooth crash. This inference is not corroborated by Ottaviani and Porcelli,[19] who give numerical and analytical evidence that electron inertia, by itself, can lead to reconnection rates that are faster than exponential (quasiexplosive) in time.

Although, by the account given above, there is disagreement at a fundamental level in Refs. 17–19, it is widely appreciated that at the high temperatures typical of present-day tokamaks, it is necessary to go beyond the basic model for **R** considered in Refs. 17–19. The four-field, two-fluid model of Hazeltine *et al.*[20] is an attempt to generalize the well-known reduced magnetohydrodynamics (hereafter RMHD) model,[21] incorporating physical effects associated with finite-ion-Larmor radius (FLR), compressibility, and electron adiabaticity. Aydemir[22] has found by numerical integration of the four-field, two-fluid equations that parallel electron pressure gradients cause rapid acceleration in the nonlinear growth of the $m=1$ mode. (Similar numerical results have also been reported by Kleva *et al.*[23]) Zakharov *et al.*[24] have argued, based on analytical approximations of the four-field model, that in the presence of a parallel electron pressure gradient, the early nonlinear growth of $m=1$ is exponential in time.

Wang and Bhattacharjee[25] have recently developed an analytical model that reproduces the salient features of $m=1$ island growth seen in Aydemir's simulation. The calculation given in Ref. 25 is based on a generalized Ohm's law that includes the effect of electron pressure gradient, but it does not rely on the assumption of isothermality, inherent in the four-field model. Recognizing that the electron pressure gradient is the key feature for accelerated growth in Aydemir's simulation, Wang and Bhattacharjee keep this effect in the generalized Ohm's law and show that it causes near-explosive growth of the $m=1$ instability. These analytical results provide a possible explanation for the sudden onset of the sawtooth collapse seen in tokamaks like JET.[16]

In this paper, we develop further the theory of the $m=1$ instability given in Ref. 25 in the context of a "modified magnetohydrodynamic" (hereafter MMHD) model, introduced in Sec. II. The basic goal of this model is to provide a framework for the inclusion of some important physical effects generally not considered in resistive MHD. The effects that we propose to study are primarily finite electron inertia and parallel electron pressure gradient. (The effect of hyperresistivity is described in the Appendix.) We make the simplifying approximation that the plasma is of constant density and hence incompressible, and furthermore, that the electron pressure p_e is a scalar. With these approximations, the quantity **R** on the right-hand side of Ohm's law in MMHD is written as

$$\mathbf{R}=\eta\mathbf{J}+\frac{4\pi}{\omega_{pe}^2}\frac{D\mathbf{J}}{Dt}-\frac{\nabla p_e}{ne}+\frac{\mathbf{J}\times\mathbf{B}}{nec}, \tag{2}$$

where n is the plasma number density and e is the magnitude of the electron charge. In Sec. II, we compare reduced equations derived from MMHD with the four-field model. In Sec. III, we use the constraint of helicity conservation[26,27] to revisit the geometrical structure of current sheets[8,12] in resistive MHD. In Sec. IV, we investigate the effect of finite electron inertia and parallel pressure gradients, and compare our results with those given in Refs. 18, 19, and 24. In Sec. V, we compare the predictions of theory with experimental results; this comparison includes not only the ubiquitous observations of fast crash and sudden onset, but also the long-standing, and, as yet unresolved, observation of incomplete current relaxation, for which we propose a possible qualitative solution. We conclude in Sec. VI with a summary and a discussion of the implications of our work.

II. THE MODIFIED MAGNETOHYDRODYNAMIC MODEL (MMHD)

A. The basic equations

The basic equations of MMHD are the same as the standard equations of incompressible resistive MHD, except for the generalized Ohm's law. The equations that are the same in the two models are the momentum equation,

$$\rho\frac{D\mathbf{u}}{Dt}=-\nabla p+\frac{\mathbf{J}\times\mathbf{B}}{c},\quad \rho=\text{const}, \tag{3}$$

the incompressibility condition,

$$\nabla\cdot\mathbf{u}=0, \tag{4}$$

and Maxwell's equations,

$$\nabla\times\mathbf{B}=\frac{4\pi}{c}\mathbf{J}, \tag{5}$$

$$\nabla\times\mathbf{E}=-\frac{1}{c}\frac{\partial\mathbf{B}}{\partial t}. \tag{6}$$

The generalized Ohm's law is written as

$$\mathbf{E}+\frac{\mathbf{u}\times\mathbf{B}}{c}=\eta\mathbf{J}+\frac{4\pi}{\omega_{pe}^2}\frac{D\mathbf{J}}{Dt}-\frac{\nabla p_e}{ne}+\frac{\mathbf{J}\times\mathbf{B}}{nec}. \tag{7}$$

Equation (7) is a special case of the generalized Ohm's law found in textbooks from which some terms have been dropped; these are terms that are multiplied by the mass ratio m_e/m_i, as well as terms such as $\nabla\cdot(\mathbf{Ju})$ that will turn out to be small once the ordering in the reconnection layer is imposed. We have also made the simplifying assumption that the electron pressure is scalar, and that ion temperature gradients can be neglected. (In particular, one can make the

simple approximation that the ions are cold.) Since n is a constant, we can write $\nabla p = \nabla p_e$ in Eq. (7). Alternatively, Eq. (7) can also be written as

$$\mathbf{E} + \frac{\mathbf{v}_e \times \mathbf{B}}{c} = \frac{4\pi}{\omega_{pe}^2} \frac{D\mathbf{J}}{Dt} - \frac{\nabla p}{ne}, \tag{8}$$

where $\mathbf{v}_e$ is the electron fluid velocity. In Ref. 25, the last term in Eq. (8) was referred to as the "Hall term," but in standard parlance, the last term in Eq. (7) bears this name.

B. Reduction of the basic equations

We simplify Eqs. (3)–(7) using the RMHD ordering.[21] We write $\mathbf{B} = B_0 \hat{\zeta} + \mathbf{B}_P$, where B_0 is the constant toroidal field, ζ is the toroidal angle, and $B_P \ll B_0$. The auxiliary magnetic field $\mathbf{B}_*$ is given by the relation

$$\mathbf{B}_* = \mathbf{B}_P - \frac{B_0}{q_s R} r \hat{\theta} = \nabla \psi \times \hat{\zeta}, \tag{9}$$

where r is the radial, θ is the poloidal, and $\zeta (= z/R)$ is the axial coordinate in a cylinder of periodicity length $2\pi R$, $q_s = 1$ at the singular surface $r = r_s$ for the $m = 1$, $n = 1$ instability, and ψ is the helical flux function. The toroidal current is given by

$$J_{\text{tor}} = \frac{c}{4\pi} \left(J + \frac{2B_0}{q_s R} \right), \tag{10}$$

where $J \equiv -\nabla^2 \psi$. Using the incompressibility condition (4) and ordering the toroidal flow to be small, we write

$$\mathbf{u} = \nabla \phi \times \hat{\zeta}, \tag{11}$$

which gives the axial vorticity

$$\omega = -\nabla^2 \phi \hat{\zeta} \equiv \omega \hat{\zeta}. \tag{12}$$

Taking the projection of (7) parallel to $\mathbf{B}$, we obtain, in RMHD ordering,

$$\frac{D\psi}{Dt} = d_e^2 \frac{DJ}{Dt} - \frac{c}{neB_0} \mathbf{B} \cdot \nabla p. \tag{13}$$

The curl of the momentum equation yields

$$4\pi\rho \frac{D\omega}{Dt} = \mathbf{B} \cdot \nabla J. \tag{14}$$

To summarize, all the fluid variables can be calculated in terms of the three scalar functions: ψ, ϕ, and p. Equations (13) and (14) are the dynamical equations for ψ and ϕ, respectively. An equation for p can be obtained, in principle, by taking the divergence of the momentum equation (3). The results that we describe here depend strongly on the presence of pressure gradients in Ohm's law, but they are not sensitively dependent on the equation of state. We demonstrate in Sec. IV that in the reconnection layer, it is convenient to obtain information about p from the projection of the momentum equation parallel to $\mathbf{B}$.

We now compare the reduced MMHD equations with the four-field equations.[20] In the four-field model, there are four scalar fields: the parallel vorticity, the flux, the electron pressure, and the parallel fluid velocity. Reduced MMHD is simpler, in that it involves the first three scalar fields of the four-field system, but not the last. This also means that reduced MMHD omits compressibility effects, which we do not consider, qualitatively, to be a serious omission for the treatment of nonlinear tearing modes in tokamaks. In comparing the two models, it should be observed that the parallel (electron) pressure gradient has different origins in the two systems. In MMHD, since the density is constant, the pressure gradient is entirely due to temperature gradients. In contrast, since the four-field model assumes that the plasma is isothermal, the pressure gradient is entirely due to variations in density. (Indeed, the equation of state in the four-field model is nothing but the equation of continuity for the electron density.) In other words, if the density is taken to be constant, there is no pressure gradient in the four-field model. This explains why Aydemir[22] attributes the novel behavior exhibited by the $m = 1$ mode in his simulation to "the coupling between the electron pressure gradient term in Ohm's law... and the parallel divergence of the electron velocity...," with the latter contributing to the variations in the density.

The inclusion of (electron) temperature gradients in MMHD has significant physical implications for the linear as well as nonlinear regimes of the $m = 1$ mode. In the linear regime, it is well known that temperature gradients have a stabilizing effect on the $m = 1$ instability.[28–31] Zakharov *et al.*[24] have presented some evidence that the occurrence of sawtooth oscillations in TFTR correlates well with the predictions of the linear instability criterion. Since we are concerned in this paper with the nonlinear evolution of the instability, we shall assume that the linear instability criterion is satisfied, and that the destabilizing effect of magnetic shear overcomes the stabilizing effect of temperature gradients.

The inclusion of temperature gradients has significant implications for the nonlinear evolution of the $m = 1$ instability. In particular, we note that the development of a current sheet in the early nonlinear stages of the instability causes a jump in the temperature across the separatrix, due to the requirement of force balance. Thus, the current sheet produces a thermal barrier that can sustain sharp differences in temperature between the inside and the outside of an $m = 1$ island in the early nonlinear stage of the instability. We discuss in Sec. V that as the island grows to a large size, this thermal barrier is dissolved on the branch of the separatrix separating the island and the hot core, and a rapid flattening of the temperature profile in the sawtooth region occurs before the reconnection is completed.

III. NONLINEAR $m = 1$ INSTABILITY IN THE RMHD MODEL

In this section, we revisit some aspects of the nonlinear theory of resistive $m = 1$ kink-tearing modes. In resistive MHD, Ohm's law (12) takes the form

$$\frac{\partial \psi}{\partial t} + \mathbf{u} \cdot \nabla \psi = \frac{\eta c^2}{4\pi} J. \tag{15}$$

Following Rosenbluth *et al.*[32] and Waelbroeck,[8] the nonlinear evolution of $m = 1$ islands can be described by a sequence of neighboring equilibria, which obey the relation

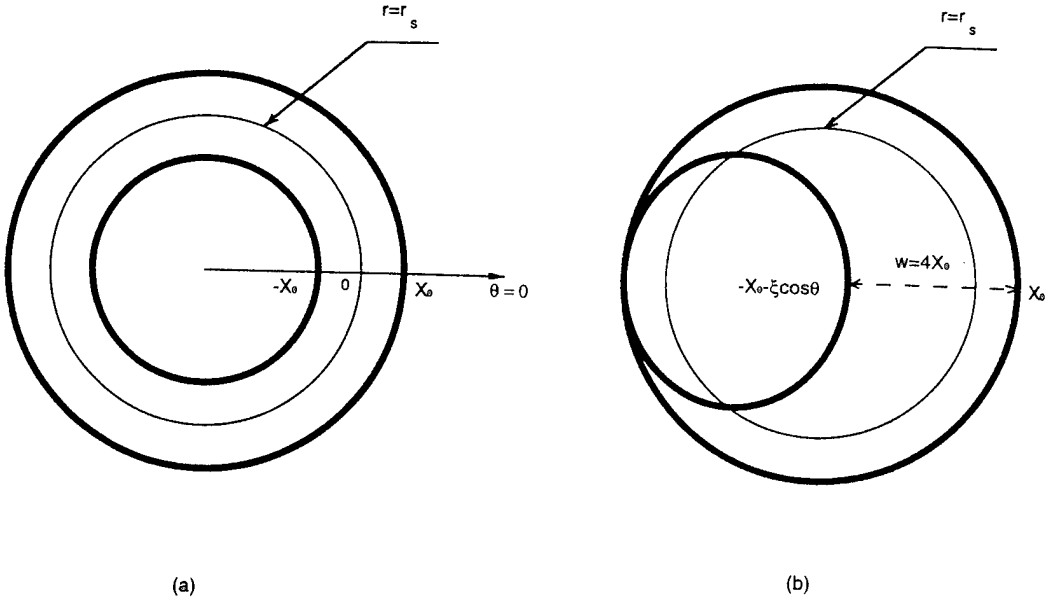

FIG. 1. The reconnection geometry for the $m=1$ mode: (a) initial state and (b) intermediate state after reconnection begins.

$$\psi(x,\vartheta)=\psi_0(x_0)=\psi_0[x+\xi(x,\vartheta)], \tag{16}$$

where ψ_0 is the flux function in the initial cylindrical equilibrium, ξ is the radial displacement, $x\equiv r-r_s$, and $\vartheta\equiv\theta-\zeta$. [Here (r,θ,ζ) is the standard cylindrical coordinate system. Our notation is somewhat different from that of Ref. 8, in that we label flux surfaces by x_0, which denotes their radial position in the unperturbed configuration, whereas x is simply the radial distance from the singular surface in the unperturbed configuration.] In the vicinity of the rational surface, we have

$$\psi_0(x)\cong\tfrac{1}{2}\psi_0''x^2, \tag{17}$$

where we have set $\psi_0(0)=0$, and

$$\psi_0''\equiv r_s(\mathbf{k}\cdot\mathbf{B})'_{r=r_s}=-B_P(r_s)q'(r_s). \tag{18}$$

Then, using (16), we can write the flux function as

$$\psi\cong\tfrac{1}{2}\psi_0''[x+\xi(x)\cos\vartheta]^2. \tag{19}$$

The nulls of $\mathbf{B}_*$ occur where $\nabla\psi=0$, which yields an O point at $\vartheta=0$ and an X point at $\vartheta=\pi$.

Figure 1 shows two flux surfaces (labeled by $x_0=\pm X_0$) that reconnect to form an island of width $w=4X_0$. For an approximate evaluation of (19), we use, in the exterior region, the linear eigenfunction[28,31,9,12]

$$\xi=\begin{cases}\xi_0, & x_0<-X_0,\\ 0, & x_0>X_0.\end{cases} \tag{20}$$

The outer and inner branches of the separatrix are given, respectively, by

$$x=X_0=w/4 \tag{21}$$

and

$$x=-\xi_0\cos\vartheta-X_0=-\xi_0\cos\vartheta-w/4. \tag{22}$$

The width of the island is

$$w=2\xi_0. \tag{23}$$

We note that the geometrical structure given by (22) and (23) differs significantly from that given in Ref. 12. It is assumed in Ref. 12 that the inner branch of the separatrix is a shifted half-circle of the same radius as the outer branch, and that the outer branch is the circle $x=0$. Furthermore, Ref. 12 argues that the inner branch coincides with the outer branch for $|\vartheta|>\pi/2$, leading to the formation of a current sheet on the latter part of the separatrix. This picture is somewhat oversimplified, because it omits the important role of the helicity conservation constraint in producing a current sheet, that not only spans the Y points, but the entire separatrix.[8]

In Refs. 26 and 27, it is shown that the $m=1$ kink-tearing mode conserves magnetic helicity contained between two reconnecting flux tubes. Two tubes with the same value of helical flux reconnect at the separatrix to form one tube. For an incompressible plasma, the constraint of helicity conservation then implies that the area enclosed between the two flux tubes before reconnection should be equal to the area of the resultant tube after reconnection. We obtain

$$\pi(r_s+X_0)^2-\pi(r_s-X_0)^2=\pi\left(r_s+\frac{w}{4}\right)^2-\int\int_{\text{inner}} r\,dr\,d\vartheta, \tag{24}$$

where

$$\int\int_{\text{inner}} r\,dr\,d\vartheta = \frac{1}{2}\int_0^{2\pi} d\vartheta\left(r_s - \frac{w}{4} - \frac{w}{2}\cos\vartheta\right)$$

$$= \pi\left(r_s - \frac{w}{4}\right)^2 + \frac{\pi}{8}w^2. \tag{25}$$

From (24) and (25), we obtain

$$X_0 = \frac{w}{4} - \frac{w^2}{32 r_s}. \tag{26a}$$

Equations (21) and (26a) motivate the definitions

$$X_i \equiv \frac{w}{4} - \frac{w^2}{32 r_s} < \frac{w}{4} \equiv X_e, \tag{26b}$$

where X_e is the innermost flux surface of the exterior (unreconnected) region and X_i is the outermost flux surface of the interior (island) region. All currents, initially located on the surfaces in the region $X_i < |x_0| < X_e$, are squeezed into a narrow strip that lies along the separatrix. Along this separatrix, the flux function remains continuous, but its normal derivative is not. Thus, a current sheet is formed, encompassing the Y points *as well as the rest of the separatrix*. The importance of this current sheet, spanning the entire separatrix, cannot be overstated, because it explains why a thermal barrier appears to separate the inside and outside of the magnetic island in observations of the early nonlinear stage of a sawtooth.

The geometry of the current sheet, sustained by the helicity-conserving reconnection process can be determined by the functions $f(x)$ and $g(\vartheta)$, defined in Ref. 8. These functions satisfy the relation

$$\frac{1}{2}\left(\frac{\partial\psi}{\partial r}\right)^2 = F(\psi) - G(\vartheta) \equiv \frac{1}{2}(\psi_0'')^2[f(\psi) - g(\vartheta)]. \tag{27}$$

In the linear or early nonlinear stage of the instability, when the island is thin, Eq. (19) reduces to

$$\psi \cong \tfrac{1}{2}\psi_0'' x^2 + x\psi_0''\xi(x)\cos\vartheta. \tag{28}$$

For such a *thin* island, near the X point, Eq. (28) gives

$$f(x_0) = x_0^2 \tag{29}$$

and

$$g(\vartheta) = \xi_0^2\cos^2\vartheta. \tag{30}$$

Near the O point, we have $g(\vartheta) = 0$. The flux surface x_0 satisfying $f(x_0) \leq g(\vartheta_m)$ will be "trapped" between the "mirror points" $\pm\vartheta_m$. Therefore, for the separatrix $x_0 \cong \pm w/4$, the "mirror points" are determined by the relation

$$\left(\frac{w}{4}\right)^2 = \left(\frac{w}{2}\right)^2\cos^2\vartheta_m, \tag{31}$$

which gives

$$\vartheta_m = \pm 2\pi/3, \tag{32}$$

as found in Ref. 24 by numerical solution of the integral equation[32] governing the current sheet. (To be precise, Ref. 24 finds numerically that $\pi - \vartheta_m = \pm 59.7°$. As demonstrated here, this angle can also be determined by simple geometrical considerations.)

In RMHD, it can be shown[8,12] that $w \sim \eta t^2$. The slowing down of an exponentially growing linear instability to algebraic (quadratic) growth in time can be attributed to the presence of the current sheet on the separatrix encompassing the Y points. This slowing-down effect is, in a way, consistent with the results of Ref. 9, where it is demonstrated that if one neglects the current sheet, then the nonlinear $m = 1$ instability grows exponentially in time.

As mentioned in the Introduction, the slow algebraic growth of the $m = 1$ islands cannot account for the sudden onset of the sawtooth crash.[16] Waelbroeck[33] has recently proposed a mechanism for the onset of the crash, based on a transitional property of $m = 1$ islands in the RMHD model. The onset occurs when a slow $m = 1$ tearing mode, exhibiting Rutherford growth and no current sheet, makes a transition to a $m = 1$ kink-tearing mode with a current sheet. Two points should be made about this suggestion. First, as Waelbroeck himself notes, the transition requires the presence of large precursor islands (of the order of 10 cm for JET), which are not observed in many sawtooth discharges. Second, the transition occurs from a regime in which the island grows as t to a regime in which the island grows as t^2, but *both* of these are algebraic regimes, much too gradual to account for the temporal suddenness seen in the data.

IV. NONLINEAR $m=1$ INSTABILITY IN MMHD

The parallel component of the generalized Ohm's law in MMHD is given by Eq. (13). In the absence of the electron pressure gradient effect in Ohm's law, the singular layer in the early reconnection stage is of the order of the collisionless skin depth d_e.[17–19] The presence of the electron pressure gradient broadens the singular layer to a width of the order of the ion-sound Larmor radius $\rho_s (\equiv c_s/\Omega_i$, where c_s is the sound speed and $\Omega_i \equiv eB_T/m_i c$ is the ion cyclotron frequency).[24] We now give a simplified treatment of nonlinear $m = 1$ dynamics in MMHD, extending our earlier results[25] on this subject. In what follows, we neglect the collisional dissipation term in Ohm's law. (The small but finite resistivity in a high-temperature plasma causes irreversible heating and reconnection, but is assumed to be subdominant to the other collisionless mechanisms that control the growth of the nonlinear instability.)

A. Near-explosive tendency for thin islands

From the incompressibility condition, we obtain

$$v_0\Delta = \int_0^{\vartheta_0} u_r r_s\,d\vartheta' = u_0 r_s\int_0^{\vartheta_0}\cos\vartheta'\,d\vartheta'$$

$$= u_0 r_s\sin\vartheta_0, \tag{33}$$

where Δ is the width of the flow channel, $\vartheta' \equiv \pi - \vartheta = \pm\vartheta_0$ designate the angular locations of the two tips of a thin $m = 1$ island, $v_0 = u_\vartheta(\vartheta_0)$, and $u_0 = u_r(\vartheta' = 0)$. As mentioned above, the presence of the pressure gradient term in Ohm's law causes a separation of small scales, i.e., $\Delta \gg \Delta_j \sim d_e$, where Δ_j is the width of the current channel.

From flux conservation in the outer region, we have $d\psi/dt = u_r B_{*\vartheta}$. Using Ohm's law (13) to match the inner limit of the outer region with the outer limit of the inner region, we obtain

$$u_r B_{*\vartheta} \cong u_r B_{*\vartheta} \frac{d_e}{\Delta} - \frac{cB_{*\vartheta}}{neB_0 r_s} \frac{\partial p}{\partial \vartheta}. \tag{34}$$

An equation for the pressure can be obtained in this case from the component of the momentum equation parallel to **B**, given by

$$\rho \mathbf{B} \cdot \frac{\partial \mathbf{u}}{\partial t} + \mathbf{B} \cdot \nabla \left(p + \frac{1}{2} \rho u^2 \right) = \rho \mathbf{B} \cdot \mathbf{u} \times \nabla \times \mathbf{u}. \tag{35}$$

Since inside the singular layer, the dominant component of **u** is $u_\vartheta(\vartheta)$ along the sheet, the right-hand side of (35) vanishes to leading order. The first term on the left-hand side is much smaller than the second term, as the plasma satisfies the force-balance condition along the sheet. Thus, we obtain[12,25]

$$\frac{\partial}{\partial \vartheta} \left(p + \frac{1}{2} \rho u_\vartheta^2 \right) \cong 0. \tag{36}$$

Equation (36) can be shown[12] to yield the outflow speed $v_0 = B_{*\vartheta}/(4\pi n m_i)^{1/2}$. Since $B_{*\vartheta} = |\psi_0''(r_s)| w/2 = B_P(r_s) w/(2r_s)$, we obtain[25]

$$v_0 = (w/2r_s) v_A, \tag{37}$$

where $v_A \equiv B_P(r_s)/(4\pi n m_i)^{1/2}$. From (23), it follows that $u_0 = dX_0/dt = \frac{1}{4} dw/dt$. Integrating Eq. (34) with respect to ϑ, and making use of Eqs. (33) and (36), we obtain

$$\frac{dX}{dt} = \frac{2 d_e \omega_A}{r_s \sin \vartheta_0} X + \frac{\omega_A^2}{2\Omega_i \sin \vartheta_0} X^2, \tag{38}$$

where we define $X \equiv w/(2r_s)$, and keep terms up to $O(X^2)$ for $X \ll 1$. The angle ϑ_0 is determined by setting $d(\vartheta_0) = \Delta$, where

$$d(\vartheta) = \frac{w}{4} - \left(-\frac{w}{4} - \xi \cos \vartheta \right) = w \sin^2 \frac{\vartheta}{2}, \tag{39}$$

is the distance between the inner and outer branches of the separatrix. Thus, we obtain

$$\sin^2 \frac{\vartheta_0}{2} = \frac{\Delta}{w}. \tag{40}$$

From (33) and (37), we obtain, for thin islands,[25]

$$\sin \vartheta_0 \cong (\omega_A/\Omega_i)^{1/2}, \tag{41}$$

where $\omega_A \equiv v_A/r_s$. Integrating Eq. (38), we obtain

$$X(t) = \frac{X(0) \exp(2\gamma_c t)}{1 - [\gamma_0/(2\gamma_c)] X(0) [\exp(2\gamma_c t) - 1]}, \tag{42}$$

where

$$\gamma_0 = \frac{\omega_A^2}{2\Omega_i \sin \vartheta_0} \tag{43}$$

and

$$\gamma_c = \frac{d_e}{r_s} \frac{\omega_A}{\sin \vartheta_0}. \tag{44}$$

It is clear that the thin-island evolution equation (42) exhibits two phases. The first phase describes exponential growth in time, given by $X(t) \approx X(0) \exp(2\gamma_c t)$, and dominated by electron inertia. In the second phase, the island growth exhibits a near-explosive tendency, since the expression on the right side of (42) has a finite-time singularity at

$$t_c = \frac{1}{2\gamma_c} \ln \left(\frac{\gamma_0}{2\gamma_c X(0)} + 1 \right). \tag{45}$$

We emphasize that the finite-time blowup suggested by (42) is merely a tendency, and is not realized in practice. This is because Eq. (42) is derived under the thin-island approximation $w/2r_s \ll 1$, which breaks down when the island size becomes a significant fraction of the plasma radius. In the next section, we shall discuss how the island passes from the near-explosive phase to the decay phase.

Equation (42) is more general than the thin-island description given in Ref. 25. It reduces to the result of Ref. 25 in the asymptotic limit $2\gamma_c(t_c - t) \gg 1$, whereby we obtain the result that $X(t)$ for thin islands tends to blow up as $(t_c - t)^{-1}$. As stated in Ref. 25, the contribution to the reconnection velocity due to the electron inertia scales as $u_c \sim (d_e/\Delta) u_r$. Since, in the presence of the electron pressure gradient term in Ohm's law, we have $\Delta \sim \rho_s$, the electron inertia term makes a smaller contribution to the reconnection velocity than the electron pressure gradient which eventually dominates the dynamics.

In the absence of the electron pressure gradient term, we have $\Delta \sim \Delta_j \sim d_e$, and electron inertia essentially controls the reconnection dynamics. In this special case, our model predicts that the thin island will tend to grow exponentially in time in the early nonlinear phase, which agrees with one of the findings of Ottaviani and Porcelli.[19] However, Ref. 19 also claims that electron inertia, by itself, can subsequently lead to near-explosive growth of islands, which is not what we find. As discussed in Ref. 19, the reduced equations without the electron pressure gradient term in Ohm's law admit the conservation of the canonical momentum $\psi - d_e^2 J$ (where $J \equiv \nabla^2 \psi$), convected by the flow. It can then be seen by simple scaling arguments[19] that the shrinking of the singular layer width Δ and the growth of a near-singular current J is a consequence of the conservation of canonical momentum. We differ from Ref. 19, however, in its conclusion that this behavior leads to near-explosive growth in the island size. (In our case, the near-explosive growth is caused by electron pressure gradients, and not by electron inertia.) In the absence of electron pressure, we expect that the current sheet will grow increasingly singular and rapidly slow down the reconnection process (in the same qualitative way as it does in RMHD) after a short-lived exponential phase. This is indeed what is shown by the numerical results of Drake and Kleva,[18] as well as by Aydemir.[22] Furthermore, Drake and Kleva note that the rapidly shrinking current layer is unstable to current-driven instabilities that generate a form of hyperresistivity.[34–37] In the Appendix, we calculate some relevant scalings of island growth in the presence of hyperre-

sistivity, and compare them with numerical results. For two possible choices of hyperresistivity, we show that the island growth is either algebraic or exponential in time.

In the presence of electron pressure gradients, the island exhibits near-explosive growth. Our conclusion, in this respect, differs from that of Zakharov *et al.*,[24] who predict exponential growth of the island using the four-field model equations. The discrepancy can be resolved by reconsidering Eq. (29) of Ref. 24, which can be written as $dw/dt \sim w\Delta$. Zakharov *et al.* suggest that Δ is time independent, and hence w grows exponentially with time. In contrast, we show in Ref. 25 as well as here that Δ itself scales as w, leading to near-explosive growth in time.

B. Large-island dynamics

The slowing down of the nearly explosive growth of the island occurs due to the reduction in the magnitude of $B_{*\vartheta}$ from the value $B_{*\vartheta} = B_P w/(2r_s)$.[25] This reduction can be calculated heuristically, using the relation

$$B_{*\vartheta} = \frac{\partial \psi_*}{\partial r} = \left. \frac{\partial \psi_0}{\partial r} \right|_{r-\xi} . \tag{46}$$

Expanding the right-hand term in (46), and keeping the first significant correction term in the expression for $B_{*\vartheta}$, we obtain

$$B_{*\vartheta} \cong |\psi_0''| \frac{\xi}{r_s}(r_s - \xi) = B_P(r_s) X(1-X). \tag{47}$$

For large islands, Eq. (41) is then modified to

$$\sin^2 \frac{\theta_0}{2} = \frac{\Delta}{w} = \frac{\omega_A}{4\Omega_i}(1-X). \tag{48}$$

Using (47) and (48), Eq. (38) can be written more generally as

$$\frac{dX}{d\tau} \cong \frac{1}{2}\left(\frac{\omega_A}{\Omega_i}\right)^{1/2} X^2(1-X)^{3/2} + \frac{2d_e}{r_s} X\left(\frac{\Omega_i}{\omega_A}(1-X)\right)^{1/2}, \tag{49}$$

where $\tau \equiv \omega_A t$. We now define the nonlinear growth rate,

$$\gamma_N \equiv \frac{d \ln X}{d\tau} . \tag{50}$$

Equation (49) is integrated with the initial condition $X(0) = \rho_i/(2r_s)$ and the JET parameters cited in Refs. 16, 17, and 25. The nonlinear growth rate γ_N can then be computed according to (50). Figure 2 shows the time history of the island growth rate, from its early exponential and near-explosive growth, followed by the decay phase. This picture shows a strong qualitative resemblance to the numerical results of Aydemir[22] [Fig. 2(b)] based on the four-field system.

V. COMPARISON WITH EXPERIMENTS

Thanks to sophisticated diagnostic development over the last decade, much is now known and observable regarding sawtooth behavior in tokamaks. These observations and the enhanced database pose a substantial challenge for any theoretical model. In this section we compare the predictions of the present model with salient features of observations.

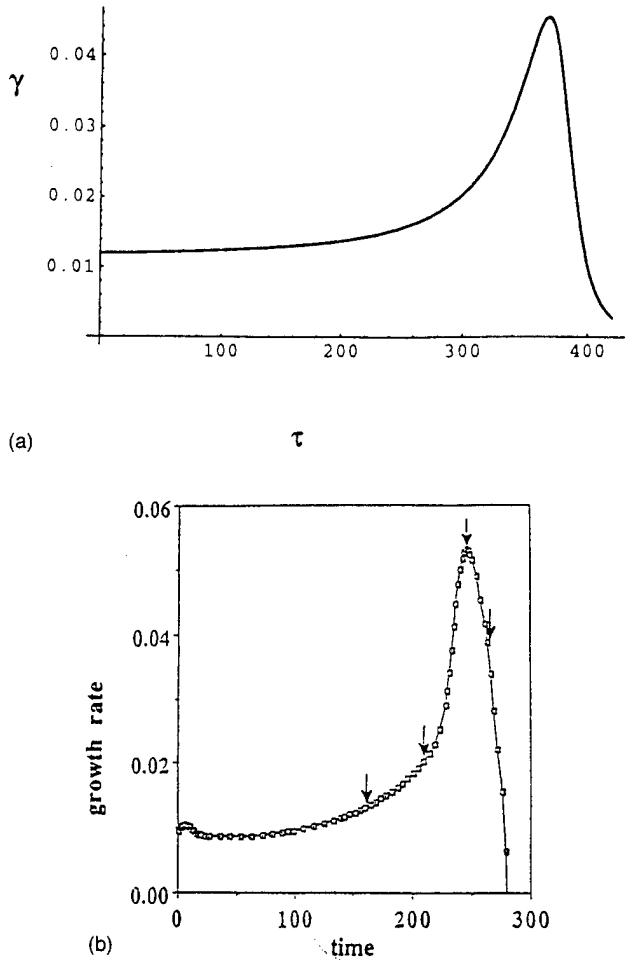

FIG. 2. (a) The nonlinear growth rate for the island width predicted by the analytical model. (b) The nonlinear growth rate obtained numerically in Ref. 22. The physical parameters and initial conditions for the numerical calculation are somewhat different than that used for (a).

A. Rapidity and onset of the sawtooth crash

In the Introduction, we have discussed that the time scale of sawtooth crashes in present-day tokamaks such as JET and TFTR is one to two orders of magnitude faster than the prediction of the Kadomtsev model. A correct theory of sawtooth crashes should account not only for this fast crash, but also for the time development of the growth rate, leading to the crash. The latter is often referred to as the "onset" or the "fast-trigger" problem,[38,39] and the issue here is the mechanism by which a sudden transition occurs from the sawtooth ramp phase to the collapse phase. It is of some interest to note that the sudden onset of sawteeth has been a persistent feature of not just the data from present-day tokamaks, but earlier tokamaks as well. Jahns *et al.*[40] point out this feature and attribute it to a transition from a linear growth phase, scaling as $\exp(\gamma_L t)$, to an even faster growth phase, scaling as $\exp(\gamma_4 t^5)$. The constant γ_4 is related in Ref. 40 to the magnetic shear, and the time dependence of the process is ascribed to the gradual time evolution of the *linear* growth

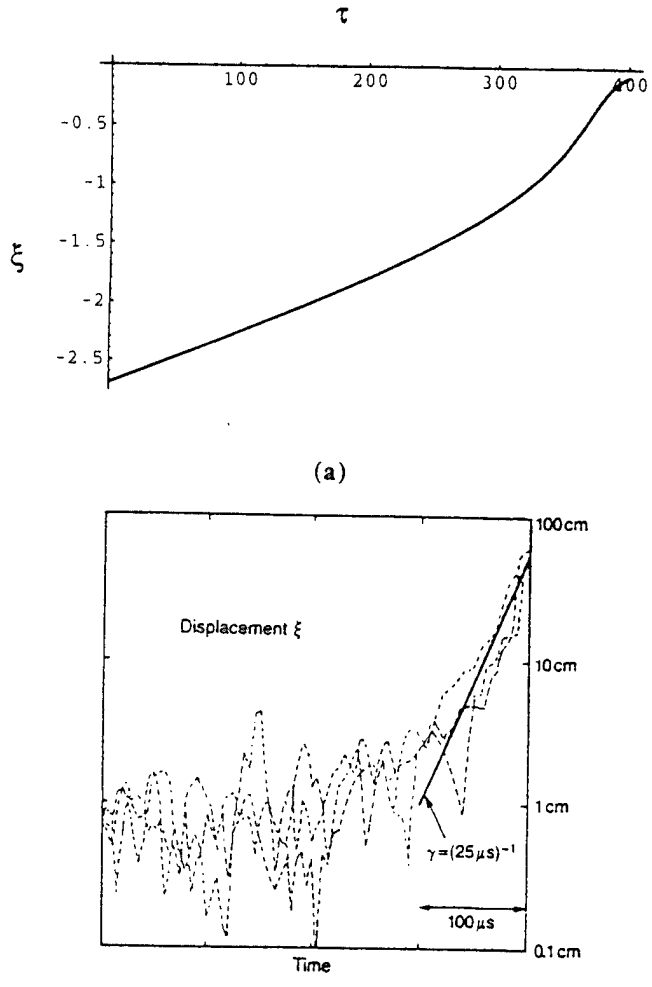

FIG. 3. (a) The plasma displacement ξ predicted by the analytical model. (b) The displacement as inferred from the peak soft-x-ray emissivity in the JET device reported in Ref. 16.

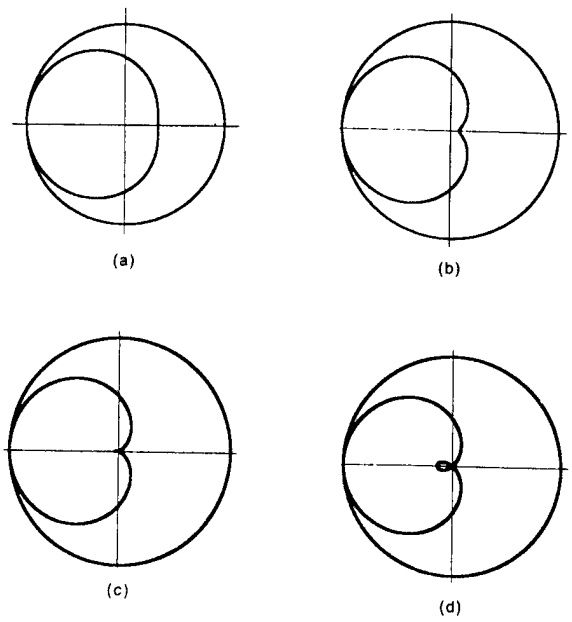

FIG. 4. Geometrical deformation of the island as the size of the island increases from (a)–(d). The shapes (c) and (d) are not physically realizable.

rate due to the temporal change in the q profile. We now know, primarily due to the theoretical developments in the last five years, the mode becomes nonlinear rather quickly, at which point its growth in the RMHD model is algebraic in time, superseding the $\exp(\gamma_4 t^5)$ phase considered in Ref. 40.

Figure 3(a) is a plot of the (logarithm of the) displacement $\xi(=Xr_s)$ versus time, obtained by integrating Eq. (49) with the initial condition $X(0)=\rho_i/(2r_s)$. The slope of the curve in the near-explosive phase is approximately $(20\ \mu s)^{-1}$, which should be compared with the slope $(25\ \mu s)^{-1}$ of the JET data, taken from Ref. 16 and shown in Fig. 3(b). The JET data also suggests that the near-explosive phase lasts approximately 100 μs. The theoretical model predicts that this time scale should be of the order of $100/\omega_A$, but since ω_A depends on the shear q' at the rational surface, the quantitative estimates may vary over a wide range (20–100 μs), depending on the value of q' at the rational surface for the $m=1$, $n=1$ mode. In fact, this is qualitatively consistent with the observation in TFTR[41] that sawtooth crash times can vary over a wide range for current profiles that look globally rather similar (but may have local dissimilarities in the value of q' at the rational surface).

It should be noted that the sudden onset to the near-explosive phase in the present model occurs from the thin-island phase of the dynamics, and therefore, does not require the presence of precursor islands. We point out that in Fig. 2(a), we follow the time evolution of the island until $\xi\sim 0.6r_s$. After this point, as is seen by inspection of Fig. 2, the nonlinear growth rate is very slow. Furthermore, as discussed in the next section, the strong island deformation induces a rapid temperature relaxation that outpaces the current relaxation due to reconnection.

B. Incomplete reconnection or the "$q(0)$ problem"

One of the outstanding problems of sawtooth research has been the experimental observation that $q(0)$, the safety factor on axis, remains fixed at a value smaller than unity ($\sim$0.7–0.8) throughout the sawtooth cycle. This was first reported by Osborne *et al.*[42] in the small tokamak TOKAPOLE II, but appears not to have been widely noticed. Results from TEXTOR by Soltwisch *et al.*,[43] and since then in other tokamaks as well, have brought this issue into sharp focus. These observations pose a major problem for theory, because the Kadomtsev model and its variants appear to lead to a complete relaxation of the temperature *and* the current profile inside the $q=1$ surface, with $q(0)$ rising above unity at the end of the crash. Lichtenberg *et al.*[44] have given a possible mechanism for partial reconnection based on the stochastization of field lines caused by secondary island formation.

Based on the present model, we propose an alternative explanation. In Fig. 4, we show the inner and outer bounding

surfaces of the island as it grows in size. The equation of the outer surface, which is circular, is given by (23), whereas the inner surface is described by (24). Figures 4(a)–4(d) correspond to increasing values of w. We note the strong distortions from near-circularity in these plots. (In the mathematical literature, the inner curve is referred to as the Limacon of Pascal, but we prefer to describe it as a "deformed crescent.") Distortions such as the ones in Figs. 4(a) and 4(b) are, in fact, seen in contour plots of the tomographically reconstructed x-ray emission (see, for instance, Fig. 4 of Ref. 14). As the island size increases rapidly due to the effect of electron pressure gradient on the $m=1$ instability, the hot core shrinks. This is also illustrated in the electron-temperature isotherms measured during a sawtooth crash in TFTR by Nagayama and co-workers,[45] described in a recent article by Callen *et al.*[46] (see Fig. 1 of Ref. 46).

What mechanism sustains the sharp temperature difference between the hot core and the island in the early stages of the island growth? It is the current sheet that lies all along the separatrix of the magnetic island described by Figs. 4(a)–4(b). That the sheet automatically sustains a steep temperature jump ΔT_e between the colder inside of the island and the outside, which comprises the hot core, can be easily seen from the pressure-balance condition $JB_{*\vartheta}/c \sim n \nabla_\perp T_e \sim n \Delta T_e/\Delta$. Once the current sheet is formed, it acts as a thermal barrier between the inside of the island and the hot core in the early stages.

As the island grows near-explosively, the inner surface becomes a deformed crescent. However, the strong distortions depicted in Figs. 4(c) and 4(d), which correspond to $\xi \geq 0.6 r_s$ and involve unphysical cusps, merely reflect a tendency, and are not realized in practice. What actually occurs is that the current sheet is drastically weakened locally in the neighborhood of where the cusp tends to form. (It is easy to see that the direction of the current due to the cusp opposes the direction of the original current sheet.) The weakening, and eventual dissolution, of the current sheet in the inner edge of the island destroys the local force balance. Once the force balance is destroyed, nonequilibrium conditions produce locally a perpendicular convection speed $u_\perp$ of the order of the sound speed, i.e., $u_\perp \sim c_s \sim \rho_s \Omega_i$. The sharp temperature jump, ΔT_e, can no longer be sustained, and the large convection velocity sets up a mixing process that begins at the boundary separating the hot core from the interior of the island, and penetrates rapidly into the island. On the other hand, the hot core shrinks, as its outer layers are peeled away, and the heat is redistributed within the island, which acts as a sink. The relevant time scales can be elucidated by considering heuristically a model temperature equation,

$$\frac{\partial T_e}{\partial t} = -\mathbf{u}\cdot\nabla T_e + \kappa_\perp \nabla_\perp^2 T_e + \kappa_\parallel \nabla_\parallel^2 T_e . \tag{51}$$

Of all the terms on the right of (51), the time scale associated with the last term is generally the shortest, because the parallel heat conduction coefficient is several orders of magnitude higher than the perpendicular heat conduction coefficient. The large parallel heat conduction produces rapid equilibration of temperature on any given flux surface. Hence, we can take the temperature to be a flux function, i.e., $T_e = T_e(\psi,t)$, and eliminate the last term by flux-surface averaging. Now, using the numbers in Ref. 25, we see that

$$\mathbf{u}\cdot\nabla T_e \approx u_\perp \frac{T_e}{\Delta} \approx \frac{\rho_s}{\Delta}\Omega_i T_e \gg \kappa_\perp \nabla_\perp^2 T_e \approx \frac{\kappa_\perp}{\Delta^2} T_e , \tag{52}$$

where we have taken $\kappa_\perp \approx 10^{-7} r_s^2 \omega_A$, based on a thermal diffusivity of 1 m^2/s.[39] Equation (52) demonstrates that the mixing of temperature across the separatrix due to nonequilibrium convection occurs much faster than that due to perpendicular heat conduction. It is important to recognize here that once the current sheet on the inner branch of the separatrix, which acts as the thermal barrier between the inside of the island and the hot core, is destroyed, the mixing caused by the strong temperature gradients ΔT_e through the fissure in the thermal barrier can no longer be held back by the surfaces away from the separatrix (on which the current distribution is smooth). Every layer of the hot core that is peeled away by this process brings the growing island in contact with another layer of a smaller hot core, and the temperature flattening between the different layers resembles a falling stack of dominos. The entire process occurs on time scales of 10–100 μs for tokamaks such as JET and TFTR. [Since the temperature relaxation in the core and the island is dominated by the mixing of the plasma by convection, the time for the mixing to be completed can be estimated roughly as $\tau_m < r_s/c_s \sim r_s/(\rho_s\Omega_i) \sim 100/\Omega_i \sim 10$ μs.]

By momentum conservation, the rapid outflow from the hot core into the island should cause a recoil of the hot core, and this can possibly account for the bulge that is seen in the fourth and fifth frames of the electron temperature isotherms in Fig. 1 of Ref. 46.

With the heat escaping from the hot core, which cools down rapidly, the parallel pressure gradient near the X point is lost before the reconnection is completed. In other words, the parallel pressure gradient that drives the near-explosive growth of a large island ends up depleting itself, because the current sheet on the inner separatrix, which acts as the thermal barrier between the core and the island, is destroyed, and the temperature equilibrates rapidly. Thus, in the scenario described above, the temperature relaxation occurs much faster than the current relaxation, with the central q value remaining below unity.

A natural question is what sustains the temperature gradient between the sawtooth region and the plasma just outside the sawtooth region. It is important to remember that while the inner surface of the island deforms due to the rapid growth in the island size, the outer surface remains circular, and still carries a current sheet that decays in time slowly. The heat pulse that "spills over" this outer thermal barrier from a sawtooth relaxation process is a weaker dynamical process than that caused by the nonequilibrium condition generated by the local dissolution of the current sheet on the inner separatrix.

It is interesting to consider why the numerical simulation of Aydemir[22] based on the four-field model, which provided the stimulus for the present work, should exhibit complete reconnection. We recall that a significant distinction between MMHD and the four-field model is that the former includes electron temperature gradients, while the latter does not. It is,

therefore, not surprising that while it captures the near-explosive growth of the instability, it cannot describe the self-consistent evolution of the temperature.[47,48] On the other hand, the self-consistent temperature evolution *is* included in the resistive MHD simulations of Aydemir *et al.*[39] In the simulations of Ref. 39 (which do not include the electron pressure gradient term in Ohm's law), the temperature relaxation seems to occur slightly ahead of the current relaxation, despite the fact that the Lundquist number for the simulation is about two orders of magnitude smaller than for the experiments. (This is clear form a close comparison of Figs. 8 and 10 of Ref. 39.) This suggests that current relaxation in a real experiment should be lagging behind temperature relaxation in a more pronounced manner than presently seen in the resistive MHD simulations, because the Lundquist numbers used in the simulations, due to numerical constraints, are too low and overestimate the rate of reconnection.

In view of the developments presented here, we suggest that if existing MHD simulations are modified by including electron pressure gradient terms in Ohm's law, and compensate for the fact that realizable Lundquist numbers for simulations overestimate the rate of reconnection, it may be possible to account for the long-standing mystery of why temperature relaxation in experiments occur while magnetic reconnection still remains incomplete. We are led to this suggestion because electron pressure gradients, by inducing near-explosive growth in the island size, should cause the inner surface to become a deformed crescent much more rapidly than they do in standard resistive MHD simulations. This, in turn, will set in motion the rapid thermal equilibration between the hot core and the island, well before the reconnection process is completed. In other words, temperature relaxation will occur well before complete current relaxation can change $q(0)$.

The picture presented above is quite different from the mechanism proposed by Kolesnichenko *et al.*[49] who have given an explanation of the $q(0)$ problem based on a scenario of secondary reconnection following the reconnection phase described by Kadomtsev. In this mechanism, the final state does have $q(0)$ less than unity, but it is preceded by an intermediate stage [Figs. 2(b) and 2(c) of Ref. 49] in which $q(0)$ rises to a value greater than unity. We note that the evolution of the temperature profile is not considered in Ref. 49. Though the scenario of secondary reconnection is plausible in principle, it may be of academic interest in practice because the rapid temperature relaxation that is observed to end a sawtooth cycle, is likely to supersede the secondary reconnection process.

VI. SUMMARY AND DISCUSSIONS

In this paper, we have developed a theory of the nonlinear collisionless $m=1$ kink-tearing instability, including the effects of electron inertia and electron pressure gradient. Three dynamical phases are identified in the growth and decay of $m=1$ islands. Due to the effect of electron inertia, the mode grows exponentially with time in the early stage of its nonlinear dynamics. However, due to the effect of the parallel electron pressure gradient, which broadens the current layer to a width of the order of the ion-sound Larmor radius, this exponential phase is followed by a near-explosive phase in which the island tends to blow up as $(t_c-t)^{-1}$. This near-explosive tendency is arrested as the island grows to a size of the order of the sawtooth region, and is followed by a rapid decay phase. The master equation (49) describes all three phases.

Viewed in its entirety, the present theoretical model provides an explanation of the salient features of sawtooth oscillations in tokamaks. The model accounts for both the rapidity of the sawtooth crash and its sudden onset, and detailed comparison with data (Sec. V) from JET shows satisfactory agreement with the observed time scales. Based on considerations of the geometrical deformations that occur in the island separatrix as the island grows near-explosively to large size, we have offered a qualitative mechanism for the shrinking of the hot core and the temperature equilibration between the hot core and the colder island. This temperature equilibration occurs rapidly, within a time scale of 10–100 μs, and well before the reconnection process can be completed. Thus, $q(0)$ remains unaffected during the sawtooth cycle.

These results have significant implications, not only for tokamaks, but also for astrophysical plasma phenomena, such as solar flares. Observations indicate three distinct phases of energy release from a solar flare: a growth phase characterized by ultraviolet (UV) and soft x-ray emission, a sudden impulsive phase characterized by UV and hard x rays, and a decay phase dominated by soft x rays.[50] It is well known that the x-ray emissions are produced by accelerated energetic particles, and that there is a substantial amount of modeling involved in relating these emissions to the signatures of MHD processes, such as the electric field produced during reconnection. The connection with the hard x-ray emission, which is dominantly a signature of prompt, high-energy electrons is perhaps the least difficult to establish, and we assume that this emission reflects qualitatively the time dependence of the electric field. We then suggest that collisionless magnetic reconnection, influenced by electron pressure gradients, is a possible candidate for the time development of the hard x-ray emission during a flare. This will be the subject of a future publication.

ACKNOWLEDGMENTS

We are grateful to Professor J. D. Callen, Dr. C. C. Hegna, Dr. M. Ottaviani, and Professor F. Porcelli for stimulating discussions.

This work is supported by National Science Foundation Grant No. ATM 93-10157 and Air Force Office of Scientific Research Grant No. F49620-93-1-0071.

APPENDIX: EFFECT OF HYPERRESISTIVITY

In this appendix, we shall discuss the effect of hyperresistivity[34–37] on the nonlinear evolution of the $m=1$ instability under conditions when it is the dominant term in the generalized Ohm's law. We then write

$$\frac{d\psi}{dt}=\mu_e d_e^2\,\nabla_\perp^2 J=\mu_e d_e^2\,\nabla_\perp^4\psi\approx\frac{r_s^4}{\tau_\mu}\frac{\Delta B_*}{\Delta_\mu^3}, \tag{A1}$$

where μ_e is a constant, $\tau_\mu \equiv r_s^4/(\mu_e d_e^2)$ is a characteristic time scale, and Δ_μ is the singular layer width in the presence of hyperresistivity. Using Eqs. (33) (with Δ replaced by Δ_μ) and (37), we obtain

$$\frac{dX}{dt} = \frac{1}{2}(\sin\vartheta_0)^{-3/4}\frac{X^{3/4}}{\tau_c}, \tag{A2}$$

where $\tau_c \equiv \tau_\mu^{1/4}\tau_A^{3/4}$ is the characteristic reconnection time, introduced in Ref. 51. Integrating Eq. (A2), we obtain

$$X = \frac{1}{2^{1/2}\sin^3\vartheta_0}\left(\frac{t}{\tau_c}\right)^4, \tag{A3}$$

which is faster than nonlinear resistive growth, but still algebraic in time.

We now consider the case in which μ_e itself is proportional to J, as in Ref. 18. In this case, Eq. (A1) changes to

$$\frac{d\psi}{dt} = d_e^2\nabla_\perp\cdot\mu_e\cdot\nabla_\perp J \approx d_e^2\frac{\hat{\beta}^{1/2}\Delta B_*}{\Delta_\mu}\left(\frac{d_e}{\Delta_\mu}\right)^3\frac{J}{(4\pi\rho)^{1/2}}$$

$$\approx \hat{\beta}^{1/2}\left(\frac{d_e}{\Delta_\mu}\right)^5\frac{(\Delta B_*)^2}{(4\pi\rho)^{1/2}}, \tag{A4}$$

where $\hat{\beta} \equiv \beta m_i/m_e$. Then, we obtain

$$\frac{dX}{dt} \approx 2\hat{\beta}^{1/2}\left(\frac{d_e}{r_s}\right)^{5/6}\left(\frac{\omega_A}{\sin\vartheta_0}\right)^{5/6}X, \tag{A5}$$

which yields exponential growth of the island, given by

$$X = X(0)\exp(2\gamma_c t), \tag{A6}$$

where

$$\gamma_c = \hat{\beta}^{1/2}\left(\frac{d_e}{r_s}\frac{\omega_A}{\sin\vartheta_0}\right)^{5/6} \tag{A7}$$

is the characteristic time scale defined by Eq. (7) of Ref. 18.

[1] P. H. Rutherford, Phys. Fluids **16**, 1903 (1973).
[2] B. V. Waddell, M. N. Rosenbluth, D. A. Monticello, and R. B. White, Nucl. Fusion **16**, 528 (1976).
[3] A. Sykes and J. A. Wesson, Phys. Rev. Lett. **37**, 140 (1976).
[4] W. Park, D. A. Monticello, and R. B. White, Phys. Fluids **27**, 137 (1984).
[5] D. Biskamp, Phys. Fluids **29**, 1520 (1986).
[6] A. Y. Aydemir, Phys. Rev. Lett. **59**, 649 (1987).
[7] L. A. Charlton, B. A. Carreras, J. A. Holmes, and V. E. Lynch, Phys. Fluids **31**, 347 (1988).
[8] F. L. Waelbroeck, Phys. Fluids. B **1**, 2372 (1989).
[9] R. D. Hazeltine, J. D. Meiss, and P. J. Morrison, Phys. Fluids **29**, 1633 (1986).
[10] P. A. Sweet, *Electromagnetic Phenomena in Cosmical Physics* (Cambridge University Press, New York, 1958), p. 123; E. N. Parker, J. Geophys. Res. **62**, 509 (1957).
[11] B. B. Kadomtsev, Fiz. Plasmy **1**, 710 (1975) [Sov. J. Plasma Phys. **1**, 389 (1975)].
[12] D. Biskamp, Phys. Fluids B **3**, 3353 (1991).
[13] X. Wang and A. Bhattacharjee, Phys. Fluids B **4**, 1795 (1992).
[14] A. W. Edwards, D. J. Campbell, W. W. Engelhardt, H.-U. Fahrbach, R. D. Gill, R. S. Granetz, S. Tsuji, B. J. D. Tubbing, A. Weller, J. Wesson, and D. Zasche, Phys. Rev. Lett. **57**, 210 (1986).
[15] K. McGuire and the TFTR group, Phys. Fluids B **2**, 1287 (1990).
[16] J. A. Wesson, A. W. Edwards, and R. S. Granetz, Nucl. Fusion **31**, 111 (1991).
[17] J. A. Wesson, Nucl. Fusion **30**, 2545 (1990).
[18] J. F. Drake and R. G. Kleva, Phys. Rev. Lett. **66**, 1458 (1991).
[19] M. Ottaviani and F. Porcelli, Phys. Rev. Lett. **71**, 3802 (1993).
[20] R. D. Hazeltine, C. T. Hsu, and P. J. Morrison, Phys. Fluids **30**, 3204 (1987).
[21] M. N. Rosenbluth, D. A. Monticello, H. R. Strauss, and R. B. White, Phys. Fluids **19**, 198 (1986); H. R. Strauss, *ibid.* **20**, 1354 (1977).
[22] A. Y. Aydemir, Phys. Fluids B **4**, 3469 (1992).
[23] R. G. Kleva, J. F. Drake, and F. L. Waelbroeck (private communication, 1992).
[24] L. E. Zakharov, B. Rogers, and S. Migliulo, Phys. Fluids B **5**, 2498 (1993). For a recent discussion of the observational aspects of sawtooth stabilization due to temperature-gradient effects, see F. M. Levinton, L. Zakharov, S. H. Batha, and M. C. Zarnstorff, Phys. Rev. Lett. **72**, 2895 (1994).
[25] X. Wang and A. Bhattacharjee, Phys. Rev. Lett. **70**, 1627 (1993).
[26] B. B. Kadomtsev, in *Plasma Physics and Controlled Nuclear Fusion Research 1976* (International Atomic Energy Agency, Vienna, 1977), Vol. I, p. 555.
[27] A. Bhattacharjee, R. L. Dewar, and D. A. Monticello, Phys. Rev. Lett. **45**, 347, 1217 (1980).
[28] G. Ara, B. Basu, B. Coppi, G. Laval, M. N. Rosenbluth, and B. V. Waddell, Ann. Phys. **112**, 443 (1978).
[29] F. Porcelli, Phys. Rev. Lett. **66**, 425 (1991).
[30] H. L. Berk, S. M. Mahajan, and Y. Z. Zhang, Phys. Fluids B **3**, 351 (1991).
[31] B. Coppi and P. Detragiache, Phys. Lett. A **168**, 59 (1992).
[32] M. N. Rosenbluth, R. Y. Dagazian, and P. H. Rutherford, Phys. Fluids **16**, 1894 (1973).
[33] F. L. Waelbroeck, Phys. Rev. Lett. **70**, 3259 (1993).
[34] P. K. Kaw, E. J. Valeo, and P. H. Rutherford, Phys. Rev. Lett. **43**, 1398 (1979).
[35] A. H. Boozer, J. Plasma Phys. **35**, 133 (1986).
[36] A. Bhattacharjee and E. Hameiri, Phys. Rev. Lett. **57**, 206 (1986).
[37] H. R. Strauss, Phys. Fluids **29**, 3668 (1986).
[38] J. A. Wesson, P. Kirby, and M. F. Nave, in *Plasma Physics and Controlled Nuclear Fusion Research 1986* (International Atomic Energy Agency, Vienna, 1987), Vol. II, p. 3.
[39] A. Y. Aydemir, J. C. Wiley, and D. W. Ross, Phys. Fluids B **1**, 774 (1989).
[40] G. L. Jahns, M. Soler, B. V. Waddell, J. D. Callen, and H. R. Hicks, Nucl. Fusion **18**, 609 (1978).
[41] K. McGuire (private communication).
[42] T. H. Osborne, R. N. Dexter, and S. C. Prager, Phys. Rev. Lett. **49**, 734 (1982).
[43] H. Soltwisch, W. Stodiek, J. Manickam, and J. Schlüter, in Ref. 38, Vol. I, p. 263.
[44] A. J. Lichtenberg, K. Itoh, S.-I. Itoh, and A. Fukuyama, Nucl. Fusion **32**, 495 (1992).
[45] Y. Nagayama, K. McGuire, and A. Cavallo, Phys. Rev. Lett. **67**, 1082 (1991).
[46] J. D. Callen, B. A. Carreras, and R. D. Stambaugh, Phys. Today **45**, 34 (1992).
[47] R. E. Denton, J. F. Drake, R. G. Kleva, and D. A. Boyd, Phys. Rev. Lett. **56**, 2447 (1986).
[48] G. Vlad and A. Bondeson, Nucl. Fusion **29**, 1139 (1989).
[49] Ya. I. Kolesnichenko, Yu. V. Yakovenko, D. Anderson, M. Lisak, and F. Wising, Phys. Rev. Lett. **68**, 3881 (1992).
[50] See, for instance, T. Sakao, T. Kosugi, S. Masuda, M. Inda, K. Makishima, R. C. Canfield, H. S. Hudson, T. R. Metcalf, J.-P. Wuelser, L. W. Acton, and Y. Ogawara, Publ. Astron. Soc. Jpn. **44**, L83 (1992).
[51] A. Y. Aydemir, Phys. Fluids B **2**, 2135 (1990).

Forced magnetic reconnection and the persistence of current sheets in static and rotating plasmas due to a sinusoidal boundary perturbation

Z. W. Ma, Xiaogang Wang, and A. Bhattacharjee
Department of Physics and Astronomy, The University of Iowa, Iowa City, Iowa 52242

(Received 12 December 1995; accepted 8 March 1996)

The problem of forced reconnection in static and rotating plasmas due to a sinusoidal boundary perturbation is revisited. The primary focus of this paper is on inner region dynamics, including the effects of resistivity as well as viscosity. It is shown that for high-Lundquist-number plasmas, the use of the "constant-ψ" approximation in the linear and nonlinear regimes of forced reconnection is not justified. The linear and nonlinear dynamics in the inner region are characterized by the persistence of current sheets. Explicit analytical solutions for the time dependence of the reconnected flux and current sheet density are given, and tested by numerical simulations. These results differ qualitatively from earlier analytical results on forced reconnection in static plasmas [T. S. Hahm and R. M. Kulsrud, Phys. Fluids **28**, 2412 (1985)] (except in a very restricted range of parameters) as well as rotating plasmas [R. Fitzpatrick and T. C. Hender, Phys. Fluids B **3**, 644 (1991)]. Some qualitative implications for laboratory and space plasmas are discussed. © *1996 American Institute of Physics.* [S1070-664X(96)02306-3]

I. INTRODUCTION

Magnetic reconnection can be either "free" or "forced." Free reconnection is caused by the spontaneous occurrence of a nonideal instability, whereas forced reconnection is caused by externally imposed boundary perturbations on a stable equilibrium.

For many systems of physical interest, both types of reconnection can coexist. To differentiate between them, we need to consider typical time scales. Let us consider an external perturbation, characterized by a time scale $\tau_0 \equiv a/V_0$, where a is the system size and V_0 is the characteristic fluid velocity involving the boundary perturbation. The magnetic field in the system responds to the external perturbation on an Alfvén time scale $\tau_A \equiv a/V_A$, where V_A is the characteristic Alfvén speed. The relevant hierarchy of time scales for forced reconnection is $\tau_A \ll \tau_0 \ll \tau_R$, where τ_R is the characteristic resistive diffusion time. The inequality $\tau_A \ll \tau_0$ allows us to assume that the plasma in the "outer region," which obeys the ideal magnetohydrodynamic (MHD) equations, is quasistatic. The presence of a small but finite resistivity causes reconnection in the "inner region" on a time scale τ_S that is much faster than the diffusion time scale τ_R. [In resistive MHD theory, τ_S is equal to $\tau_L \equiv \tau_A^{2/3}\tau_R^{1/3}$ in a linear non-constant-ψ regime,[1] to $\tau_{cL} \equiv \tau_A^{2/5}\tau_R^{3/5}$ in a linear constant-ψ regime,[2] and to $\tau_{SP} \equiv (\tau_A \tau_R)^{1/2}$ in a nonlinear non-constant-ψ regime.[3–5]] Hence, we obtain $\tau_A \ll \tau_S \ll \tau_R$. Since τ_0 and τ_S are both bounded by τ_A and τ_R, we should consider the relative magnitudes of τ_0 and τ_S. If reconnection occurs when $\tau_S \ll \tau_0$, it does so because of the free energy already stored in the equilibrium, and is hence "free." On the other hand, if $\tau_S \gg \tau_0$, reconnection is "forced." In this paper, we focus on forced reconnection, for which the relevant hierarchy of time scales is $\tau_A \ll \tau_0 \ll \tau_S \ll \tau_R$.

Hahm and Kulsrud (hereafter HK)[6] obtained linear as well as nonlinear analytical solutions for a paradigmatic model of forced reconnection in static plasmas, proposed by Taylor. Wang and Bhattacharjee[7] (hereafter WB) demonstrated that the nonlinear treatment of HK is valid only in a very restricted parameter range for high-Lundquist-number plasmas because it ignored the persistence of current sheets in a "non-constant-ψ" regime.[8,9] Fitzpatrick and Hender[10] (hereafter FH) extended HKs results by including the effects of rotation and viscosity in cylindrical geometry, but did not address the issue of persistence of current sheets in a nonlinear, non-constant-ψ regime. Later, Fitzpatrick *et al.*[11] did touch upon this issue briefly, but maintained that the reconnection layer probably "remains constant-ψ throughout its linear (and subsequent nonlinear) phase."

The main goal of this paper is to demonstrate that the use of the constant-ψ approximation in the linear and nonlinear regimes of forced reconnection is incorrect, both with and without the effect of plasma rotation and viscosity. In the regime of non-constant-ψ regime, current sheets do persist, and the time evolution of the island during forced reconnection is qualitatively different than that obtained from the standard Rutherford theory[12] and its generalizations.[6,10,11]

The following is a layout of this paper. In Sec. II, we introduce the basic model and dynamical equations. In Sec. III, we summarize the analytical results obtained in Refs. 5 and 7 for a stationary plasma and give numerical results from a two-dimensional (2-D) resistive MHD simulation that corroborate the analytical scalings. In the following sections, we present analytical and numerical results on forced reconnection in a resistive and viscous stationary plasma (Sec. IV), as well as a resistive and viscous rotating plasma (Sec. V). In all of these cases, we demonstrate the persistence of a nonlinear non-constant-ψ regime with a current sheet. We conclude in Sec. VI with qualitative discussions on the implications of our results.

II. MODEL AND EQUATIONS

We use a Cartesian coordinate system. For configurations with translation symmetry along z, we represent the magnetic field $\mathbf{B}$ as

 1070-664X/96/3(6)/2427/7/$10.00 © 1996 American Institute of Physics

$$\mathbf{B}=\hat{\mathbf{z}}\times\nabla\psi+B_z\hat{\mathbf{z}}, \tag{1}$$

where ψ is a flux function and B_z is constant. An incompressible flow field $\mathbf{v}$ with the same symmetry can also be represented as

$$\mathbf{v}=\hat{\mathbf{z}}\times\nabla\phi, \tag{2}$$

where ϕ is a streamfunction and $\nu_z=0$. The incompressible MHD equations for the two dependent variables, ψ and ϕ, can then be written (in cgs units) as

$$\frac{\partial\psi}{\partial t}+(\phi,\psi)=\frac{\eta c^2}{4\pi}\nabla^2\psi, \tag{3}$$

$$\frac{\partial\omega}{\partial t}+(\phi,\omega)=\frac{1}{c\rho}(\psi,J_z)+\frac{\nu}{\rho}\nabla^2\omega, \tag{4}$$

where η is the resistivity, ν is the viscosity, c is the speed of light, $J_z=(c/4\pi)\nabla^2\psi$ is the longitudinal (z component of the) current density, $\omega=\nabla^2\phi$ is the z component of the vorticity, ρ is the (constant) density of the plasma, and $(f,g)=\hat{\mathbf{z}}\cdot\nabla f\times\nabla g$. The initial static equilibrium is given by

$$\psi_0=aB_0\ \ln[\cosh(x/a)], \tag{5}$$

where B_0 and a are constants.

We cast Eqs. (3), (4), and (5) in dimensionless form by redefining $\mathbf{B}/B_0\rightarrow\mathbf{B}$, $\mathbf{x}/a\rightarrow\mathbf{x}$, $t/\tau_A\rightarrow t$, $\psi\rightarrow\psi/B_0a$, $\phi\rightarrow\phi/V_Aa$, $\omega\rightarrow\omega a/V_A$, $\mathbf{J}\rightarrow 4\pi a\mathbf{J}/(cB_0)$, where $\tau_A=a/V_A$ and $V_A=B_0/(4\pi\rho)^{1/2}$ is the Alfvén speed. The dimensionless equations are

$$\frac{\partial\psi}{\partial t}+(\phi,\psi)=\frac{1}{S}\nabla^2\psi, \tag{6}$$

$$\frac{\partial\omega}{\partial t}+(\phi,\omega)=(\psi,J_z)+\frac{1}{R}\nabla^2\omega, \tag{7}$$

$$\psi_0=\ln(\cosh x), \tag{8}$$

where $S\equiv\tau_R/\tau_A$ is the Lundquist number, with $\tau_R\equiv 4\pi a^2/\eta c^2$, and $R\equiv\tau_\nu/\tau_A$ is the Reynolds number, with $\tau_\nu\equiv\rho a^2/\nu$.

For the numerical results discussed below, we solve Eqs. (6) and (7) simultaneously using a Runge–Kutta finite-differencing scheme that has fourth-order accuracy in time and second-order accuracy in space. Exploiting the symmetry of the initial conditions, we carry out the numerical simulation in a quadrant of the physical domain in the $x-y$ plane, with $x\in(0,1)$ and $y\in(0,2)$. At the boundary $x=0$ and $y=0$, we impose symmetric boundary conditions. We assume that the boundary $y=2$ is open and impose the conditions $\partial\omega/\partial t=0$ and $\partial\psi/\partial t=-\mathbf{v}\cdot\nabla\psi$. In order to reduce the amplitude of waves reflected from the boundary, we apply a damping term at the outermost grid points. We use a three-point differencing method to calculate the first derivative at the boundary, ensuring second-order spatial accuracy. To reduce numerical error and computer time, we employ a nonuniform mesh. Nonuniform meshing enables us to increase the spatial resolution near the separatrix (where the current sheet develops).

At the lateral boundary $x=1$, the externally driven perturbation is turned on at $t=0$ and attains the asymptotic value

$$\xi_x(x=1)=\xi_0\cos(ky-\Omega t), \tag{9}$$

in a short period t_0 (typically $t_0=1$), where $\xi_0\ll 1$ is a constant and Ω is the oscillation frequency of the applied perturbation. In Taylor's model, to be discussed next, we set $\Omega=0$.

III. FORCED RECONNECTION IN TAYLOR'S MODEL

Taylor's model, first considered by HK, deals with forced reconnection in a static plasma. This model has already been treated analytically in Refs. 6 and 7, and we review the results here briefly in order to facilitate comparison with the numerical results discussed below.

A. The outer region

For $|x|\leq 1$, we approximate the initial equilibrium (8) as

$$\psi_0=x^2/2. \tag{10}$$

The equation for the perturbed boundary is

$$x=\pm(1+\xi_0\cos ky), \tag{11}$$

where we have redefined $\xi_0/a\rightarrow\xi_0$ and $ka\rightarrow k$. We write the perturbed flux function in the form

$$\tilde{\psi}(x,y,t)=\tilde{\psi}(x,t)\cos ky, \tag{12}$$

where $\tilde{\psi}(1,t)=\xi_0$. We assume that the outer region is quasi-static and obeys the ideal MHD equations. Equation (6) then reduces to

$$\frac{\partial^2\tilde{\psi}}{\partial x^2}-k^2\tilde{\psi}=0, \tag{13}$$

which has the solution[6]

$$\tilde{\psi}(x,t)=\tilde{\psi}(0,t)\left(\cosh kx-\frac{\sinh|kx|}{\tanh k}\right)+\xi_0\frac{\sinh|kx|}{\sinh k}. \tag{14}$$

For forced reconnection, at $t=0$, we have $\tilde{\psi}(0,t)=0$, and the outer region solution exhibits a jump discontinuity at $x=0$. It is clear from these simple considerations that the linear dynamics of forced reconnection begins from a non-constant-ψ regime in which $\Delta'=[\tilde{\psi}'(0+,t)-\tilde{\psi}'(0-,t)]/\tilde{\psi}(0,t)\rightarrow\infty$, where the prime (over $\tilde{\psi}$) denotes the derivative with respect to x. To determine if the non-constant-ψ regime persists for later times, we need to consider the time dependence of $\tilde{\psi}(0,t)$, obtained from the inner region solution.

B. The inner region

The linear analytical solution for the inner region is obtained by solving the resistive MHD equations.[6,7] In the linear ideal MHD phase, valid for $0<t\leq 1$, HK obtain the exact solution,[6]

$$\tilde{J}_L(0,t)=\frac{4k^2\xi_0}{\pi\sinh k}t. \tag{15}$$

In the linear resistive phase, valid for $1\leq t\ll S^{1/3}$, the reconnected flux grows in time as[6]

$$\tilde{\psi}_L(0,t)=\frac{2k^2\xi_0}{\pi\sinh k}\left(\frac{t}{\tau_{SP}}\right)^2, \tag{16}$$

where $\tau_{SP}=S^{1/2}$. The time development of the current sheet during the linear resistive phase can also be calculated by the same method and shows that the current sheet growth in this phase is still given by the ideal result (15). Hence, the "ideal" current sheet persists in the linear regime of forced reconnection.[6,7]

The transition to the nonlinear stage occurs when

$$\frac{w}{\Delta_L}\cong\frac{4k^2}{\pi}\left[\frac{2}{\pi}\frac{\xi_0}{\sinh k}\right]^{1/2}\left(\frac{t}{\tau_c}\right)^2\sim 1, \tag{17}$$

where $\tau_c=S^{1/4}$ and $\Delta_L=\pi/(2kt)$ is the linear current channel width (which is an upper bound for the reconnection layer width). Hence, the transition to the nonlinear non-constant-ψ regime occurs, superseding the linear constant ψ regime described by HK, at $t\sim\tau_c/\xi_0^{1/4}$ for modes with $k\sim 1$. However, as discussed by WB, in high-Lundquist-number plasmas ($S\geqslant 10^6$), for very short-wavelength modes ($k\gg 10$), or very long-wavelength modes ($k\ll 0.1$), the transition occurs in a constant-ψ regime. HK's analysis is valid, and the transition occurs when

$$t>\frac{\pi^{1/2}}{2k}\left(\frac{\pi\sinh k}{2\xi_0}\right)^{1/4}\tau_c\gg\tau_c. \tag{18}$$

In the nonlinear non-constant-ψ regime with a current sheet, a variant of the Sweet–Parker model holds.[7–9] The nonlinear reconnected flux is given by[7,8]

$$\tilde{\psi}_N(0,t)\cong\sqrt{2}\left(\frac{k\xi_0}{\sinh k}\right)^{3/2}\frac{t}{\tau_{SP}}. \tag{19}$$

The nonlinear current sheet saturates with an amplitude,

$$\tilde{J}_N(0,t)\cong\sqrt{2}\left(\frac{k\xi_0}{\sinh k}\right)^{3/2}S^{1/2}. \tag{20}$$

C. Numerical results

We have tested the analytical results discussed above with the 2-D, incompressible MHD code described in Sec. II. We have run simulations with Lundquist numbers (a) $S=10^5$ and (b) $S=10^4$ with $\xi_0=0.01\sqrt{\pi}$ and $k=\pi/2$. According to analytical theory, relevant time scales for these two cases are (a) $\tau_c=18$, $\tau_L=46$; and (b) $\tau_c=10$, $\tau_L=21$. Figures 1(a) and 1(b) show the time evolution of the current sheet amplitude at $x=0$ for the cases (a) and (b), respectively. In the linear regime, the analytical result predicts that the current sheet amplitude grows as $\tilde{J}_L(0,t)=J_0t$, where $J_0=0.024$ for the parameters used in the simulation. In the nonlinear stage, the total saturated current sheet amplitude, $J_{z0}+\tilde{J}_L+\tilde{J}_N$, is analytically predicted to be 2.03 for case (a) and 1.43 for case (b), respectively. We see that the simulation results (solid lines) are reproduced well by the curve $J_{z0}+J_0(t-t_0)$ in the linear phase (dot–dashed line). (Here the offset t_0 is approximately the "switch-on" time required by the perturbation to reach the maximum value ξ_0.) The numerical simulation finds that the multiplicative factor J_0 and the saturated value of the current sheet are, respectively, 0.036 and 2.5 for case (a) and 0.028 and 1.46 for case (b), in reasonably good agreement with analytical predictions. The maximum amplitude of the current sheet increases with the Lundquist number, qualitatively consistent with the analytical scalings. In Fig. 2, we show a typical three-dimensional picture of the current sheet, taken at $t=50$ for case (a).

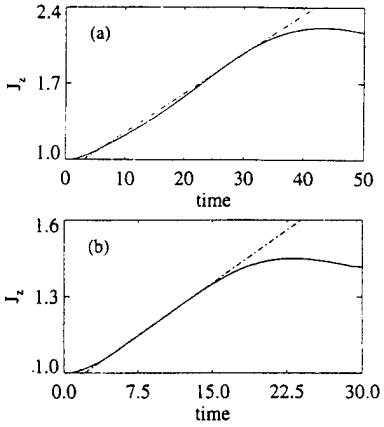

FIG. 1. The time evolution of current sheets. The solid lines represent numerical simulation results. The dashed lines are linear analytic predictions for $\xi_0=0.01\sqrt{\pi}$, $k=\pi/2$, and Lundquist numbers (a) $S=10^4$, (b) $S=10^5$.

IV. THE EFFECT OF VISCOSITY

The quasistatic outer region solution (14) is unaffected by the presence of viscosity. We consider the inner region solution, modified by viscosity.

A. The linear regime

On the ideal time scale $0<t\leqslant 1$, the exact ideal solution (8) continues to hold. The linear nonideal layer equations are

$$\frac{\partial\tilde{\psi}}{\partial t}-kx\tilde{\phi}=\frac{1}{S}\tilde{\psi}'', \tag{21a}$$

$$\frac{\partial\tilde{\phi}''}{\partial t}+kx\tilde{\psi}''=\frac{1}{R}\tilde{\phi}''''. \tag{21b}$$

We define $\partial h/\partial t\equiv -k\tilde{\phi}$, and take Laplace transforms of the time-dependent variables to write Eqs. (21) in the form

$$s(\Psi+xH)=S^{-}\Psi'', \tag{22a}$$

$$s^2H''=k^2x\Psi''+sR^{-1}H'''', \tag{22b}$$

where

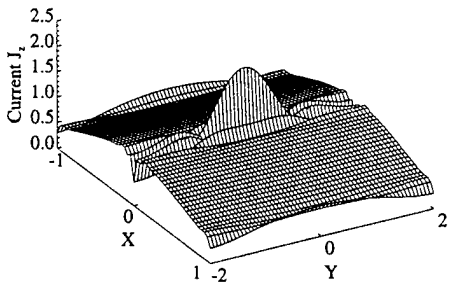

FIG. 2. A typical nonlinear current sheet in Taylor's model (at $t=50$) for case (a) of Fig. 1.

$$\Psi(x)=\int_0^\infty dt\ e^{-st}\tilde{\psi}(x,t) \tag{23a}$$

and

$$H(x)=\int_0^\infty dt\ e^{-st}h(x,t). \tag{23b}$$

There are two spatial scales in the inner region: the flow channel width that scales as $\epsilon_V=R^{-1/3}$ and the current channel width that scales as $\epsilon=S^{-1/3}$. We now consider two limiting cases.

(i) $\epsilon\gg\epsilon_V$ $(R\gg S)$

Near the neutral line $x=0$, we have $s\Psi\sim S^{-1}\Psi''$, which motivates the ordering $s\sim\epsilon\sim S^{-1/3}$. We define the variables[6] $\mu\equiv s/\epsilon$, $\theta\equiv x/\epsilon$, $Z\equiv\Psi''$, and $\Gamma\equiv\epsilon H$. Ignoring the viscous term, we write Eqs. (22) in the form

$$Z=\mu(\Psi+\theta\Gamma), \tag{24a}$$

$$\Gamma''\cong\frac{k^2\theta}{\mu^2}Z=\frac{k^2\theta}{\mu^2}(\Psi+\theta\Gamma). \tag{24b}$$

In the phase $t\ll\tau_L$ or $\mu=s\tau_L\gg1$, we obtain

$$Z\cong\mu\Psi(0)\left(1+\frac{k^2\theta^2}{\mu^2}\right)^{-2}. \tag{25}$$

We now match asymptotically the solution (25) to the outer solution (14). We note that

$$\lim_{x\to\infty}\left.\frac{d\Psi}{dx}\right|_{\text{inner}}=\int_0^\infty dx\,\frac{d^2\Psi}{dx^2}=\frac{1}{\epsilon}\int_0^\infty d\theta\,Z=\frac{\pi\mu^2}{4k\epsilon}\Psi(0) \tag{26a}$$

and

$$\lim_{x\to0}\left.\frac{d\Psi}{dx}\right|_{\text{outer}}=\frac{k\xi_0}{s\sinh k}-\frac{k\Psi(0)}{\tanh k}. \tag{26b}$$

Equating (26a) and (26b), we obtain

$$\Psi(0)\cong\frac{4}{\pi}\frac{\xi_0k^2}{s^3\tau_{SP}^2\sinh k}. \tag{27}$$

Taking the inverse Laplace transform of (27) and (25), we obtain, respectively, (16) for the reconnected flux, and (15) for the time evolution of the current sheet.

In the region $\theta\sim\epsilon_V/\epsilon\ll1$, the velocity field has a fine spatial structure. Since this fine structure does not have a significant effect on the reconnected flux or the current sheet, we do not consider it further.

(ii) $\epsilon_v\gg\epsilon(S\gg R)$

We introduce a "middle" region scale,

$$\epsilon_D\equiv(S/R)^{1/6}\epsilon, \tag{28}$$

larger than inner region scale ϵ. In the middle region, the plasma is "viscoideal," i.e., the flux remains "frozen in" but the flow is damped by viscosity. We define the variables $\mu_D\equiv s\epsilon_D^2/\epsilon^3\equiv s\tau_{LD}=sS^{2/3}/R^{1/3}$, $\theta=x/\epsilon_D$, $Z\equiv\Psi''$, and $\Gamma=\epsilon_D H$, and rewrite Eqs. (22) as

$$Z=\mu_D(\Psi+\theta\Gamma), \tag{29a}$$

$$\Gamma''=\alpha\left(\frac{k^2\theta}{\mu_D^2}\Psi''+\frac{\Gamma''''}{\mu_D}\right), \tag{29b}$$

where $\alpha\equiv S/R\gg1$. In the limit $\mu_D\gg1(t\ll\tau_{LD})$, Eqs. (29) yield to leading order,

$$\Psi+\theta\Gamma\cong0, \tag{30a}$$

$$\Gamma''\cong\frac{\alpha}{\mu_D}\Gamma''''. \tag{30b}$$

We note that the sheet current, proportional to Ψ'', is inside the channel $x\sim\epsilon\ll\epsilon_D$ and drops out of the middle region. For $\theta\gg0$, Eq. (30b) has the solution

$$\Gamma''\cong\Gamma_0\exp(-\sqrt{\mu_D/\alpha}\,\theta), \tag{31}$$

where Γ_0 is a constant. From Eq. (30a), we obtain

$$Z\cong-[2\Gamma'+\theta\Gamma'']. \tag{32}$$

Consequently, Eq. (31) yields

$$Z=\Gamma_0\exp(-\sqrt{\mu_D/\alpha}\,\theta)\left(2\sqrt{\frac{\alpha}{\mu_D}}-\theta\right). \tag{33}$$

Since

$$Z(\theta\to0)=\mu_D\Psi(\theta\to0)\cong\mu_D\Psi\left(\frac{\epsilon}{\epsilon_D}\right)=2\sqrt{\frac{\alpha}{\mu_D}}\Gamma_0, \tag{34}$$

we obtain

$$\Gamma_0=\frac{\mu_D^{3/2}\Psi(\epsilon/\epsilon_D)}{2\alpha^{1/2}}. \tag{35}$$

From Eq. (26a), it follows that

$$\Psi\left(\frac{\epsilon}{\epsilon_D}\right)\cong\Psi(0)+\Psi'(0)\frac{\epsilon}{\epsilon_D}=\Psi(0)\left(1+\frac{\pi\mu^2}{4k}\right). \tag{36}$$

Asymptotic matching between the inner, middle, and outer regions requires that

$$\lim_{x\to0}\left.\frac{d\Psi}{dx}\right|_{\text{outer}}-\lim_{x\to\infty}\left.\frac{d\Psi}{dx}\right|_{\text{inner}}=\int_0^\infty dx\left.\frac{d^2\Psi}{dx^2}\right|_{\text{middle}}. \tag{37}$$

We can write the right-hand side of Eq. (37) as

$$\frac{1}{\epsilon_D}\int_0^\infty d\theta\,Z=\frac{\mu_D^{1/2}\alpha^{1/2}\Psi(\epsilon/\epsilon_D)}{2\epsilon_D}\cong\frac{\mu_D^{1/2}\alpha^{1/2}\Psi(0)}{2\epsilon_D}\left(1+\frac{\pi\mu^2}{4k}\right). \tag{38}$$

Using Eqs. (26a), (26b), (36), and (37), we obtain for $\mu_D\gg1(t\ll\tau_{LD})$,

$$\Psi(0)\cong\frac{8}{\pi}\frac{\xi_0k^2}{\sinh k}\frac{1}{\tau_{LV}^{5/2}s^{7/2}}, \tag{39}$$

where $\tau_{LV}\equiv(\tau_V/\tau_A)^{2/15}\ \tau_{LD}=(\tau_V/\tau_A)^{2/15}\alpha^{1/3}\tau_L$. The inverse Laplace transform of Eq. (39) yields

$$\tilde{\psi}_L(0,t)\cong\frac{32\xi_0k^2}{15\pi^{3/2}\sinh k}\left(\frac{t}{\tau_{LV}}\right)^{5/2}. \tag{40}$$

The current sheet amplitude grows algebraically in time according to the equation

$$\tilde{J}_L(0,t)\cong\frac{16\xi_0 k^2\alpha^{1/2}}{3\pi^{3/2}\sinh k}\frac{R^{1/2}}{S^{2/3}}t^{3/2}. \tag{41}$$

B. The nonlinear regime

As in Sec. III B, we can demonstrate using Eqs. (40) and (41) that the transition to the nonlinear non-constant-ψ regime occurs for $t\ll\tau_{LD}$. It can be shown that

$$\frac{w}{\Delta_L}\cong\frac{16}{15}k^2\xi_0^{1/2}\left(\frac{6}{5}\pi^{9/2}\sinh k\right)^{-1/2}\left(\frac{t}{\tau_{cD}}\right)^{11/4}, \tag{42}$$

where $\tau_{cD}\equiv\tau_A^{2/11}\tau_{LD}^{9/11}\ll\tau_{LD}$, which implies that the transition to the nonlinear regime occurs when $\xi_0^{-2/11}\tau_{cD}<t$ for $k\sim1$. The constant-ψ regime is valid either for the short-wavelength modes $k\gg30$, or the long-wavelength modes $k\ll0.1$.

Following Refs. 7–9, the nonlinear reconnected flux can be shown to be

$$\tilde{\psi}_N(0,t)\cong\sqrt{2}\left[\frac{k\xi_0}{\sinh k}\right]^{3/2}\frac{t}{\tau_{SPD}}, \tag{43}$$

where $\tau_{SPD}\equiv(1+\alpha)^{1/4}\tau_{SP}$. We note that Eq. (43) is similar to Eq. (19), except for the replacement of τ_{SP} by τ_{SPD}. The saturated current sheet amplitude is given by

$$\tilde{J}_N(0,t)\cong\sqrt{2}\left(\frac{k\xi_0}{\sinh k}\right)^{3/2}S^{1/2}\left(\frac{1}{\alpha+1}\right)^{1/4}. \tag{44}$$

V. THE EFFECT OF ROTATION

Forced reconnection in a rotating viscous plasma has been considered in detail by FH with a view to applying the theory to tokamaks. In such an application, the neutral line in equilibrium (8) corresponds qualitatively to a resonant (rational) surface in a cylindrical tokamak. If the plasma has a (constant) flow velocity $V_0\hat{\mathbf{y}}$ at the neutral line, and the oscillation frequency of the applied perturbation is Ω_0, then the Doppler-shifted frequency is $\Omega=\Omega_0-kV_0$. If $\Omega=0$, then a Galilean transformation yields Taylor's model to which the considerations of Secs. III and IV apply. On the other hand, if $\Omega\neq0$, Galilean transformation to a frame moving with a velocity $V_0\hat{\mathbf{y}}$ yields a perturbed boundary,

$$x=\pm[1+\xi_0\cos(ky-\Omega t)], \tag{45}$$

considered by FH. It is assumed that $\Omega\ll1$. In other words, the Doppler-shifted frequency is assumed to be much smaller than the characteristic inverse Alfvén time. This permits us to assume that the plasma in the outer ideal region moves with the boundary. The perturbed flux at the boundary can be written as

$$\begin{aligned}\tilde{\psi}(1,y,t)&=\tilde{\psi}(1,t)\cos(ky-\Omega t)\\&\equiv\xi_0(\cos\Omega t\cos ky+\sin\Omega t\sin ky).\end{aligned} \tag{46}$$

The linearized ideal solution in the outer region can be written as

$$\tilde{\psi}(x,y,t)=\tilde{\psi}_c(x,t)\cos ky+\tilde{\psi}_s(x,t)\sin ky, \tag{47}$$

subject to the boundary conditions

$$\tilde{\psi}_c(1,t)=\xi_0\cos\Omega t,\quad \tilde{\psi}_s(1,t)=\xi_0\sin\Omega t. \tag{48}$$

A. The outer region

By Eqs. (47) and (48), the outer ideal solutions for $\Omega\leqslant1$ are

$$\begin{aligned}\tilde{\psi}_c(x,t)&=\tilde{\psi}_c(0,t)\left(\cosh kx-\frac{\sinh|kx|}{\tanh k}\right)\\&\quad+\frac{\sinh|kx|}{\sinh k}\xi_0\cos\Omega t,\end{aligned} \tag{49a}$$

$$\begin{aligned}\tilde{\psi}_s(x,t)&=\tilde{\psi}_s(0,t)\left(\cos kx-\frac{\sinh|kx|}{\tanh k}\right)\\&\quad+\frac{\sinh|kx|}{\sinh k}\xi_0\sin\Omega t.\end{aligned} \tag{49b}$$

Clearly, there is a phase shift at $x=0$, as the island attempts to catch up with the boundary perturbation. The island velocity at $x=0$, caused by the boundary rotation, can be determined from

$$\begin{aligned}\tilde{\psi}(x\to0,y,t)&=\tilde{\psi}_c(0,t)\cos ky+\tilde{\psi}_s(0,t)\sin ky,\\&\equiv\sqrt{\tilde{\psi}_c^2(0,t)+\tilde{\psi}_s^2(0,t)}\cos[ky-\vartheta(t)]\\&=\tilde{\psi}(0,t)\cos[ky-\vartheta(t)],\end{aligned} \tag{50}$$

where $\tilde{\psi}(0,t)\equiv\sqrt{\tilde{\psi}_c^2(0,t)+\tilde{\psi}_s^2(0,t)}$, $\vartheta(t)\equiv\tan^{-1}[\tilde{\psi}_s(0,t)/\tilde{\psi}_c(0,t)]$, and

$$V_w=\frac{d\vartheta(t)}{k\,dt}. \tag{51}$$

B. The linear phase

We can solve the linearized MHD equations (21) for the inner region separately for the cosine and sine components. Following the steps in Sec. IV, for $\Omega t\ll\Omega\tau_{LD}\ll1$, we obtain

$$\Psi_s(0)\cong\frac{\Omega}{s}\Psi_c(0)\cong\frac{8}{\pi}\frac{\xi_0k^2}{\sinh k}\frac{\Omega}{\tau_{LV}^{5/2}s^{9/2}},$$

which gives

$$\tilde{\psi}_c(0,t)\cong\frac{32\xi_0k^2}{15\pi^{3/2}\sinh k}\left(\frac{t}{\tau_{LV}}\right)^{5/2} \tag{52}$$

and

$$\tilde{\psi}_s(0,t)\cong\frac{2}{7}\Omega t\tilde{\psi}_c(0,t). \tag{53}$$

Therefore, by Eqs. (50) and (51), we obtain for $\Omega t\ll1$,

$$\tilde{\psi}(x\to0,y,t)=\tilde{\psi}_c(0,t)\cos\left(ky-\frac{2}{7}\Omega t\right) \tag{54}$$

and

$$V_w=\frac{2\Omega}{7k}. \tag{55}$$

Equation (55) shows that the island is dragged along by the rotating boundary.

C. The nonlinear phase

To study the transition to the nonlinear regime, we begin from the linear phase with $\Omega t \ll \Omega \tau_{LD} \ll 1$, i.e.,

$$\Omega \ll \frac{1}{\tau_{LD}} \equiv \frac{R^{1/3}}{S^{2/3}}, \tag{56}$$

which is identical to Eq. (39) of Ref. 11, where it is claimed to be a criterion for the linear layer to be constant ψ. This claim is clearly contradicted by the results presented here. Furthermore, as shown by Eq. (42), the system makes a transition to a nonlinear non-constant-ψ regime for $t \ll \tau_{LD}$. Alternatively, we can arrive at the same conclusion by noting that roughly at the transition time $\tau_{cD} \ll \tau_{LD}$, we have

$$\Delta' a \cong \frac{15\pi^{5/2}}{16k} (S^{2/3} R^{2/5})^{5/11} \gg 1, \tag{57}$$

In the nonlinear phase, following the procedure that yields Eqs. (19) and (20) [or Eqs. (43) and (44)], in the frame moving with the current sheet, we obtain

$$\frac{d\tilde{\psi}_N(0,t)}{dt} \cong \frac{\sqrt{2}}{\tau_{SPD}} [\tilde{B}_{y,\text{outer}}(x \to 0,t)]^{3/2}. \tag{58}$$

From Eqs. (49), assuming $\tilde{\psi}/\xi_0 \ll 1$, we obtain

$$\tilde{B}_{y,\text{outer}}(x \to 0,t) \cong \frac{k\xi_0}{\sinh k}. \tag{59}$$

Then, in the laboratory frame, we have

$$\frac{d\tilde{\psi}_N(0,y,t)}{dt} \cong \frac{\sqrt{2}}{\tau_{SPD}} \left(\frac{k\xi_0}{\sinh k} \right)^{3/2} \cos(ky - \Omega t). \tag{60}$$

Therefore, the oscillatory part of $\tilde{\psi}_N(0,y,t)$ is

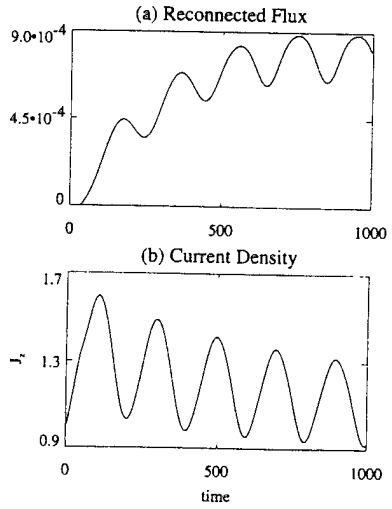

FIG. 3. The nonlinear time evolution of (a) reconnected flux and (b) current density in a viscous plasma driven by a rotating boundary perturbation. The parameters are $k=\pi/2$, $\xi_0=0.05\sqrt{\pi}$. $S=10^5$, $R=10^5$, and $\Omega=10^{-2}\pi$.

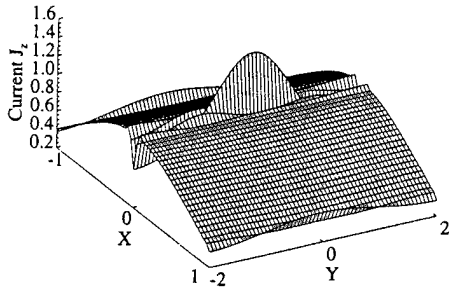

FIG. 4. A typical nonlinear current sheet in a viscous and rotating plasma (at $t=100$) for the parameters of Fig. 3.

$$\tilde{\psi}_{N,\text{osc}}(0,y,t) \cong \frac{\sqrt{2}}{\tau_{SPD}} \left(\frac{k\xi_0}{\sinh k} \right)^{3/2} \frac{[1-\sin(ky-\Omega t)]}{\Omega}. \tag{61}$$

The nonlinear current sheet is given by

$$\tilde{J}_N(0,y,t) \cong \sqrt{2} \left(\frac{k\xi_0}{\sinh k} \right)^{3/2} \frac{S^{1/2}}{(1+\alpha)^{1/4}} \cos(ky - \Omega t). \tag{62}$$

D. Numerical results

In Figs. 3(a) and 3(b), we plot the reconnected flux and the current sheet amplitude, for the parameters $k=\pi/2$, $\xi_0=5\sqrt{\pi}\times10^{-3}$, $S=10^5$, $R=10^5$, and $\Omega=10^{-2}\pi$. For these parameters, Eq. (61) yields $|\tilde{\psi}_{N,\text{osc}}|_{\max} \cong 1.2 \times 10^{-4}$, which should be compared with the numerical value $\sim 2.4\times10^{-3}$ in Fig. 3(a). For the chosen set of parameters, Eq. (62) yields $|\tilde{J}_{N,\text{osc}}|_{\max} \cong 0.35$, which should be compared with the numerical value ~ 0.4 in Fig. 3(b). In Fig. 4, we show a three-dimensional plot of a current sheet during the time evolution of the system. The current sheet is locked to the boundary perturbation in this case.

VI. CONCLUSIONS

The main goal of this paper has been to elucidate, by analysis and numerical simulation, the role of current sheets in the dynamics of forced reconnection in viscous plasmas, with or without rotation. The persistence of these current sheets invalidates the constant-ψ approximation in the linear as well as the nonlinear regimes. For magnetotail or solar coronal plasmas driven by the solar wind, these conclusions imply the persistence of current sheet structures, even if the driving electric field has a moving component. These results also suggest that previous analyses[10,11] must be reconsidered in so far as the problem of locked modes in toroidal plasmas is concerned. This will be the subject of a separate publication.

ACKNOWLEDGMENTS

We thank Dr. Chris Hegna for several useful discussions.

This research is supported by the Air Force Office of Scientific Research Grant No. F49620-93-1-10071 and the National Science Foundation Grants No. ATM 93-10157 and No. ATM 93-96280. Supercomputing resources were provided by the San Diego Supercomputing Center.

[1]B. Coppi, R. Galvao, M. N. Rosenbluth, and P. H. Rutherford, Sov. J. Plasma Phys. **2**, 533 (1976).
[2]H. P. Furth, J. Killeen, and M. N. Rosenbluth, Phys. Fluids **6**, 459 (1963).
[3]P. A. Sweet, *Electromagnetic Phenomena in Cosmical Physics* (Cambridge University Press, New York, 1958), p. 123.
[4]E. N. Parker, J. Geophys. Res. **62**, 509 (1957).
[5]B. B. Kadomtsev, Fiz. Plazmy **1**, 710 (1975) [Sov. J. Plasma Phys. **1**, 389 (1975)].
[6]T. S. Hahm and R. M. Kulsrud, Phys. Fluids **28**, 2412 (1985).
[7]X. Wang and A. Bhattacharjee, Phys. Fluids B **4**, 1795 (1992).
[8]W. Park, D. A. Monticello, and R. B. White, Phys. Fluids **27**, 137 (1984).
[9]F. L. Waelbroeck, Phys. Fluids B **1**, 2372 (1989).
[10]R. Fitzpatrick and T. C. Hender, Phys. Fluids B **3**, 644 (1991).
[11]R. Fitzpatrick, R. J. Hastie, T. J. Martin, and C. M. Roach, Nucl. Fusion **33**, 1533 (1993).
[12]P. H. Rutherford, Phys. Fluids **16**, 1903 (1973).

Forced reconnection and mode locking in rotating cylindrical plasmas

Xiaogang Wang and A. Bhattacharjee
Department of Physics and Astronomy, University of Iowa, Iowa City, Iowa 52242

(Received 10 September 1996; accepted 14 October 1996)

The problem of mode locking due to forced magnetic reconnection in rotating cylindrical plasmas is revisited. Forced reconnection is characterized by very large values of the parameter Δ', which makes the constant-ψ approximation generally inapplicable in the linear and early nonlinear regimes. The nonlinear dynamics of rotating non-constant-ψ islands is distinguished by the persistence of current sheets spanning Y points. Mode locking due to a suddenly imposed error field is discussed. Temporal dynamics and locking thresholds that differ significantly from the predictions of the constant-ψ theory [R. Fitzpatrick and T. C. Hender, Phys. Fluids B **3**, 644 (1991)] are obtained. The predictions of the present theory are compared with experimental tokamak observations. © *1997 American Institute of Physics.* [S1070-664X(97)01502-4]

I. INTRODUCTION

Mode locking has been observed in tokamaks as well as reversed-field pinches.[1–10] Locked modes caused by field errors can be very damaging to device operation. After these modes lock to the error field, they can grow further to reduce energy confinement, inhibit high-confinement mode (H-mode) operation, or even cause disruptions.

Resonant field errors can drive magnetic reconnection in a resistively stable plasma. A simple but illuminating model for this phenomenon was proposed by Taylor for static plasmas. Taylor's model was first presented and treated analytically by Hahm and Kulsrud[11] (hereafter HK) and subsequently by Wang and Bhattacharjee[12] (hereafter WB). Since the parameter Δ' is generally very large for forced reconnection, WB emphasized that the linear as well as the nonlinear island dynamics does not obey the constant-ψ approximation in most cases of physical interest. As discussed by Park *et al.*[13] and Waelbroeck,[14] the nonlinear dynamics of forced reconnection is characterized by the persistence of quasi-ideal current sheets spanning Y points. These structures are absent from the nonlinear constant-ψ islands described by the Rutherford theory.[15]

Fitzpatrick and co-workers[16–18] have extended HKs calculation to rotating viscous plasmas, relying on the constant-ψ approximation in the linear as well as nonlinear regimes. Ma, Wang, and Bhattacharjee[19] (hereafter MWB) have demonstrated recently that the constant-ψ approximation also fails in this case, and that the failure of the approximation in the linear and nonlinear regimes can be attributed to the extremely large values of Δ' characterizing forced reconnection. As in the static resistive case,[12–14] there are fundamental distinctions between constant-ψ and nonconstant-ψ islands in the presence of viscosity and plasma rotation. Nonconstant-ψ islands have current sheets and inner layer structures that are qualitatively different from constant-ψ islands, with the consequence that their temporal evolution is quite different from that of constant-ψ islands.

The main purpose of this paper is to revisit the problem of mode locking in cylindrical plasmas, considered by Fitzpatrick and Hender (hereafter FH).[16] We build on the basic results of MWB, but consider in greater detail the time history and conditions that lead to mode locking in a rotating plasma.

We discuss mode locking due to the application of an external resonant magnetic perturbation that is switched on suddenly. (This is exactly the problem discussed by HK and FH.) We calculate the relevant thresholds for the size of field errors and give a time history of the mode locking process. In cases of physical interest, we find that the parametric dependencies of the critical error fields can be significantly different from the constant-ψ theory.[16] [Theoretical projections[4] of low thresholds for such error fields have been a concern for the International Thermonuclear Experimental Reactor (ITER) device.[20]]

The layout of this paper is as follows. In Sec. II, we discuss the ideal outer region solution in a cylindrical tokamak. In Sec. III, we consider mode locking due to an error field that is switched on suddenly. In Sec. IV, we compare our analytical predictions with experimental observations from COMPASS-C,[21] and consider possible constraints on error fields for ITER. We conclude in Sec. V with a brief summary.

II. THE IDEAL OUTER REGION

Following Fitzpatrick, who has given a readable account of earlier work in Refs. 22 and 23, we consider magnetostatic equilibria in a straight periodic cylinder of radius a and length $2\pi R_0$, assumed to represent a large-aspect-ratio, low-β tokamak with circular flux surfaces. The equilibrium quantities depend only on the radial coordinate r. The components of the equilibrium magnetic field are

$$\mathbf{B}=[0,B_\theta(r),B_\varphi], \tag{1}$$

where $\varphi=z/R_0$ is an effective toroidal angle and the toroidal field B_φ is assumed to be constant. We write

$$J_\varphi=\frac{c}{4\pi r}\frac{d}{dr}(rB_\theta) \tag{2}$$

and

$$q(r)\equiv\frac{rB_\varphi}{R_0B_\theta}. \tag{3}$$

 1070-664X/97/4(3)/748/7/$10.00 © 1997 American Institute of Physics

At a rational surface $r=r_s$, we have $q(r_s)=m/n$, where m and n are integers.

The perturbed magnetic field can be written as

$$\mathbf{b}=\nabla\psi\times\hat{\mathbf{z}}. \tag{4}$$

In the ideal outer region, assuming that the plasma rotation velocity is much less than local Alfvén speed, we take the curl of the linearized momentum equation (without inertia) and obtain

$$\nabla^2\psi+\frac{4\pi J'_\varphi}{cB_\theta(nq/m-1)}\psi=0, \tag{5}$$

where the prime indicates a derivative with respect to r, and

$$\nabla^2\psi\approx\frac{1}{r}\frac{d}{dr}\left(r\frac{d\psi}{dr}\right)-\frac{m^2}{r^2}\psi, \tag{6}$$

for a large-aspect-ratio tokamak. The perturbed current density is given by

$$\frac{4\pi}{c}\mathbf{j}=\nabla\times\mathbf{b}=-\nabla^2\psi\hat{\mathbf{z}}. \tag{7}$$

If there is a tearing instability, the logarithm of the eigenfunction,

$$\psi=\psi(r,t)\ \exp\left[-i(m\theta-n\varphi)\right], \tag{8}$$

has a positive jump discontinuity at the rational surface $r=r_s$, represented by Δ'_0. On the other hand, if $\Delta'_0<0$, the tearing mode is stable. Then $\psi(r,t)\to 0$ as $t\to\infty$, and there is no free reconnection at the rational surface. However, reconnection on the rational surface can be forced by a perturbation imposed at the boundary, such as an error field. Forced reconnection is not an instability, but evolves out of the stable magnetohydrodynamic (MHD) continuum due to the influence of the perturbation. The appropriate mathematical technique to treat such an evolution (which can have algebraic dependence on time, even in the linear regime) is the initial-value method pioneered in plasma physics by Landau.

As discussed by HK and WB, forced reconnection is initiated in a nonconstant-ψ linear regime, similar to the linear regime for $m=1$ kink-tearing modes. The tearing stability parameter Δ' can be written as

$$\Delta'=\Delta'_0+\Delta'_{\rm inh},$$

where Δ'_0 is the contribution of the homogeneous solution (as in the case of free reconnection), and $\Delta'_{\rm inh}$ is the contribution of the inhomogeneous solution. Let us assume that at $t=0$, the perturbation is switched on suddenly. In the linearly perturbed state, we can write

$$\Delta'_{\rm inh}=\frac{\tilde{\psi}'(r_s^+)-\tilde{\psi}'(r_s^-)}{\tilde{\psi}(r_s)}.$$

Note that the quantity $\Delta\tilde{\psi}'(r_s)\equiv\tilde{\psi}'(r_s^+)-\tilde{\psi}'(r_s^-)$ is nonzero and measures the jump discontinuity of the helical magnetic perturbation. Since the reconnected flux is zero [i.e., $\tilde{\psi}(r_s)=0$] at $t=0$, we obtain $\Delta'_{\rm inh}\to\infty$ at $t=0$. This implies that the forced reconnection dynamics must begin in a linear nonconstant-ψ regime. HK were well aware of this fact and treated the early linear dynamics for static plasmas correctly without invoking the constant-ψ approximation, but Fitzpatrick[17] argued that the constant-ψ approximation holds for rotating viscous plasmas (in the so-called viscoresistive limit). A critique of Fitzpatrick's arguments is given in the Appendix.

As in Refs. 22 and 23, we assume that the equilibrium current J_φ is negligible *outside* the rational surface. The general solution for Eq. (5) can then be written as

$$\psi=\psi_R+\psi_D, \tag{9}$$

where ψ_R is the reconnected flux and ψ_D is the "shielded" flux,[22,23] given by

$$\psi_D=\begin{cases}\Psi_V F(r), & r_s<r<a,\\ 0, & 0<r<r_s.\end{cases} \tag{10}$$

Here Ψ_V is the vacuum flux parametrizing the field-error amplitude and the function $F(r)$ is given by

$$F(r)\equiv\left(\frac{r}{r_s}\right)^m-\left(\frac{r}{r_s}\right)^{-m}. \tag{11}$$

In general, the reconnected flux ψ_R can be written as $\psi_R=\Psi_R\psi_m(r)$, where $\Psi_R=\psi(r=r_s,t)$ is the reconnected flux at the rational surface, and $\psi_m(r)$ is a scalar function that is normalized to unity at the rational surface and satisfies the appropriate boundary conditions. In the case of locked mode eigenfunctions that penetrate to regions outside of the plasma and the vacuum vessel, the appropriate boundary condition is $\psi_m(r)\to 0$ as $r\to\infty$.

It is instructive to compare the outer solution (9) with the earlier results of HK by making some simplifying assumptions. We assume that the plasma is bounded by a perfectly conducting wall, which implies that $\psi_m(a)=0$ (assumed by HK as well as FH). Furthermore, we assume that the spatial gradients in the outer region current profile are much smaller than that in the inner region current sheet and set $J'_\varphi=0$ in the outer region (also assumed by HK as well as FH). This simplifying assumption amounts to neglecting the effect of the equilibrium current in the outer region, and can be justified on the grounds that the forced reconnection dynamics is dominated by the parameter $\Delta'_{\rm inh}$, which is controlled by the boundary perturbations. [Of course, this assumption can be relaxed and solutions can be obtained for Eq. (5) which will yield, for instance, a more accurate value for the stability parameter Δ'_0]. Under these simplifying assumptions, we obtain the following analytical solution for $\psi_m(r)$:

$$\psi_m=\begin{cases}G(r)-F(r)G(a)/F(a), & r_s<r<a,\\ G(r), & 0<r<r_s,\end{cases} \tag{12}$$

where

$$G(r)\equiv\left(\frac{r}{r_s}\right)^m. \tag{13}$$

Equations (10) and (12) yield neighboring equilibrium solutions similar to equilibria I and II given by HK, with $\psi(a)=\Psi_V F(a)\equiv\Psi_a$. The nature of these equilibrium solutions is determined by the value of Ψ_R. If $\Psi_R=0$, there is no reconnection and we obtain an equilibrium with an ideal current sheet, similar to equilibrium I of HK. If

$\Psi_R=\Psi_V F(a)/G(a)$, the reconnection flux assumes its full value. Then the equilibrium is smooth, has no ideal current sheet, and similar to equilibrium II of HK.

We will not make explicit use of the solution (12) in the rest of the paper and proceed instead with the more general from $\psi_R=\Psi_R\psi_m(r)$, obtained with less restrictive boundary conditions. Nonetheless, the discussion above makes it qualitatively clear that we have two possible neighboring equilibria of types I and II, and that the reconnection process begins in a nonconstant-ψ regime. We now proceed to obtain the time-dependent inner region solutions to determine the dynamical accessibility of these types of equilibria.

III. INNER REGION DYNAMICS

We introduce a few basic time scales: ideal, $\tau_A=r_s/v_A[v_A=B_\theta(r_s)/(4\pi\rho)^{1/2}]$, resistive, $\tau_R\equiv 4\pi r_s^2/\eta c^2$, and viscous, $\tau_V\equiv\rho r_s^2/\mu_\perp$, where ρ is the mass density, c is the speed of light, η is the parallel resistivity and $\mu_\perp$ is the perpendicular viscosity. We now consider the case of sudden error fields, treated by FH using the constant-ψ approximation. The correct inner region solution in slab geometry has been discussed by MWB. In what follows, we neglect the effect of curvature and adapt MWB's slab solution at the neutral line to the $r=r_s$ surface in a cylinder by identifying $x\equiv r-r_s$ and $k\equiv ms(r_s)/r_s$ with $s(r_s)\equiv r_s q'(r_s)/q(r_s)$.

A. The linear nonconstant-ψ regime

In a conventional tokamak, the Lundquist number $S(\equiv\tau_R/\tau_A)$ is usually much larger than the Reynolds number $R(\equiv\tau_V/\tau_A)$. In this regime, the results obtained in Secs. IV and V of MWB are relevant. Since the field error is generally fixed in position while the plasma rotates, one way to derive the relevant analytical results is by repeating the initial-value treatment, in the laboratory frame, carried out in Sec. IV of MWB. We assume that $\Omega_0\tau_A\ll 1$, where Ω_0 is the plasma rotation frequency at the $r=r_s$ surface.

We represent the perturbed velocity as $\mathbf{v}=\nabla\phi\times\hat{z}$, where

$$\phi(\mathbf{x},t)=i\phi(x,t)\ \exp[-i(m\theta-n\varphi)]. \tag{14}$$

In the inner region, it is convenient to take into account the effect of plasma rotation by making the transformations

$$\hat{\psi}(x,t)=\psi(x,t)\ \exp(-i\Omega_0 t), \tag{15a}$$

$$\hat{\phi}(x,t)=\phi(x,t)\ \exp(-i\Omega_0 t). \tag{15b}$$

Introducing the dimensionless variables $t/\tau_A\to t$, $x/r_s\to x$, $kr_s\to k$, $\hat{\psi}/r_sB_\theta(r_s)\to\hat{\psi}$, $\hat{\phi}/r_sv_A\to\hat{\phi}$, we obtain the linear nonideal layer equations:

$$\frac{\partial\hat{\psi}}{\partial t}-kx\hat{\phi}=\frac{1}{S}\psi'', \tag{16a}$$

$$\frac{\partial\phi''}{\partial t}+kx\psi''=\frac{1}{R}\hat{\phi}''''. \tag{16b}$$

Defining $\partial\hat{h}/\partial t\equiv-k\hat{\phi}$, Eqs. (16) can be rewritten as

$$\frac{\partial}{\partial t}(\hat{\psi}+x\hat{h})=\frac{1}{S}\hat{\psi}', \tag{17a}$$

$$\frac{\partial^2 h''}{\partial t^2}=k^2x\hat{\psi}''+\frac{1}{R}\frac{\partial}{\partial t}\hat{h}''''. \tag{17b}$$

Taking Laplace transforms of both sides of Eqs. (17), we obtain

$$s(\bar{\Psi}+x\bar{H})=S^{-1}\bar{\Psi}'', \tag{18a}$$

$$s^2\bar{H}''=k^2x\bar{\Psi}''+sR^{-1}\bar{H}'''', \tag{18b}$$

where

$$\bar{\Psi}(x)\equiv\int_0^\infty dt\ e^{-pt}\psi(x,t)=\int_0^\infty dt\ e^{-(p-i\Omega_0)t}\hat{\psi}(x,t), \tag{19a}$$

$$\bar{H}(x)=\int_0^\infty dt\ e^{-pt}h(x,t)=\int_0^\infty dt\ e^{-(p-i\Omega_0)t}\hat{h}(x,t), \tag{19b}$$

with $\hat{h}(x,t)=h(x,t)\ \exp(-i\Omega_0 t)$ and $s\equiv p-i\Omega_0$.

There are two spatial scales in the inner region: the flow channel width that scales as $\epsilon_V=R^{-1/3}$ and the current channel width that scales as $\epsilon=S^{-1/3}$. As mentioned above, the regime of interest for tokamaks is usually $\epsilon_V\gg\epsilon(S\gg R)$, for which we follow the inner layer solution given in Sec. IV of MWB. We introduce a "middle" region scale $\epsilon_D\equiv(S/R)^{1/6}\epsilon$, larger than inner region scale ϵ. In the middle region, the plasma is "visco-ideal," i.e., the flux remains "frozen in" but the flow is damped by viscosity. The asymptotic matching condition between the inner, middle, and outer regions is [Eq. (37) of MWB],

$$\lim_{x\to 0}\left.\frac{d\bar{\Psi}}{dx}\right|_{\text{outer}}-\lim_{x\to\infty}\left.\frac{d\bar{\Psi}}{dx}\right|_{\text{inner}}=\int_0^\infty dx\left.\frac{d^2\bar{\Psi}}{dx^2}\right|_{\text{middle}}. \tag{20}$$

By Eq. (26a) of MWB, we have

$$\lim_{x\to\infty}\left(\frac{d\bar{\Psi}}{dx}\right)_{\text{inner}}\approx\frac{\pi\mu^2}{4k\epsilon}\bar{\Psi}_R, \tag{21}$$

where $\mu=s\tau_L\equiv(p-i\Omega_0)\tau_L$, $\epsilon=\tau_L^{-1}\equiv S^{-1/3}$, and $\bar{\Psi}_R$ is the Laplace transform of Ψ_R. We also have [Eq. (38) of MWB],

$$\int_0^\infty dx\left.\frac{d^2\bar{\Psi}}{dx^2}\right|_{\text{middle}}\approx\frac{\mu_D^{1/2}\alpha^{1/2}}{2\epsilon_D}\left(1+\frac{\pi\mu^2}{4k}\right)\bar{\Psi}_R, \tag{22}$$

where $\epsilon_D\equiv(S/R)^{1/6}\epsilon$ and $\mu_D\equiv(p-i\Omega_0)\epsilon_D^2/\epsilon^3=(p-i\Omega_0)S^{2/3}/R^{1/3}\equiv(p-i\Omega_0)\tau_{LD}$. Taking the Laplace transform of the outer region solution [Eqs. (9) and (10) of Sec. II], we obtain

$$\lim_{x\to 0}\left(\frac{d\bar{\Psi}}{dx}\right)_{\text{outer}}\approx\frac{2k\Psi_V}{p}. \tag{23}$$

Substituting Eqs. (21)–(23) in the matching condition (20), and inverting the Laplace transform, we obtain for $\Omega_0 t\ll 1$,

$$\Psi_R\approx\frac{64m^2s^2(r_s)}{15\pi^{3/2}}\left(\frac{t}{\tau_{LV}}\right)^{5/2}\Psi_V\exp\left(i\frac{5}{7}\Omega_0 t\right), \tag{24}$$

where $\tau_{LV} \equiv R^{2/15}\tau_{LD}$. Equation (24) shows that the reconnected flux has a lower "effective" frequency than the plasma rotation frequency, as pointed out by MWB. Although the island rotates at a speed lower than the plasma with respect to the wall, there is no locking at this stage of the dynamics.

B. The nonlinear nonconstant-ψ regime

It has been shown by MWB that the transition to the nonlinear regime occurs generally when the island is in a linear nonconstant-ψ regime. In the nonlinear regime, we can use a variant[12–14] of the Sweet–Parker model.[24,25] Despite the apparent simplicity of the Sweet–Parker model, comparison of analytical and numerical results for static[13] as well as rotating[19] plasmas suggests that the analytical predictions are accurate to the extent of a multiplicative constant of order unity. Hence, we continue to rely on a simply variant of this model. In the reference frame moving with the current sheet, we obtain

$$\left|\frac{d\Psi_R}{dt}\right| \approx \frac{[2m\Psi_V]^{3/2}}{[r_s B_\theta(r_s)]^{1/2}\tau_{SPV}}, \tag{25}$$

where $\tau_{SPV} \equiv (\tau_A \tau_R)^{1/2}(1+\tau_R/\tau_V)^{1/4}$.

Since the current sheet rotates with the island, we obtain

$$\frac{d\Psi_R}{dt} \approx \frac{[2m\Psi_V]^{3/2}}{[r_s B_\theta(r_s)]^{1/2}\tau_{SPV}} \exp(i\Omega t), \tag{26}$$

where $\Omega = \Omega_0 - n\,\Delta\Omega_\phi$ and $\Delta\Omega_\phi$ is the change in the toroidal rotation profile at the rational surface induced by the error field (assuming that the poloidal rotation profile has been damped by viscosity). Integrating Eq. (26), we obtain

$$\Psi_R \approx -i\frac{[2m\Psi_V]^{3/2}}{\Omega\tau_{SPV}[r_s B_\theta(r_s)]^{1/2}} \exp(i\Omega t). \tag{27}$$

C. Torques in the nonconstant-ψ regime

Following Fitzpatrick,[17,22,23] we can calculate the average electromagnetic (EM) and viscous torques over the reconnection layer, but we do so in the nonconstant-ψ regime. The average toroidal EM torque is given by

$$T_{\phi EM} = \frac{\pi n}{2} R_0 \,\mathrm{Im}\left(r_s \psi^* \frac{d\psi}{dr}\right)_{r_s^-}^{r_s^+}$$

$$= \frac{\pi}{2} n R_0 \frac{[2m\Psi_V]^{5/2}}{[r_s B_\theta(r_s)]^{1/2}\Omega\tau_{SPV}}. \tag{28}$$

The viscosity of the plasma tends to relax the rotation profile to the unperturbed equilibrium state. However, in the early stages of reconnection, the relaxation is partial and occurs only in the vicinity of the reconnection layer. The corresponding localized average viscous torque can be written as

$$T_{\phi VS}^{upper} = -\frac{4\pi^2 R_0^3 r_s^2 \rho\,\Delta\Omega_{\phi s}}{\lambda_p \tau_V}, \tag{29a}$$

where $\Delta\Omega_{\phi s} \equiv \Delta\Omega_\phi(r_s)$, $\lambda_p \approx \int_{r_s^-}^{r_s^+}[\mu_\perp(r_s)dr/r\mu_\perp(r)] \approx w/r_s$, $\mu_\perp(r)$ is the perpendicular viscosity, and w is the width of the island. In contrast, if the rotation profile is fully relaxed across the entire plasma cross section, the average viscous torque becomes

$$T_{\phi VS}^{lower} = -\frac{4\pi^2 R_0^3 r_s^2 \rho\,\Delta\Omega_{\phi s}}{\lambda_e \tau_V}, \tag{29b}$$

where $\lambda_e \equiv \int_{r_s}^{a}[\mu_\perp(r_s)dr/r\mu_\perp(r)]$ is a term of order of unity. In both cases, the mode-locking condition is determined by the torque-balance equation,

$$T_{\phi EM} + T_{\phi VS} = 0, \tag{30}$$

which yields the quadratic equation

$$\Omega(\Omega - \Omega_0) + \frac{\lambda_{p,e}\tau_V}{2\tau_A^2\tau_{SPV}}\left(\frac{R_0}{r_s}\right)^{1/2}\left(\frac{b_s}{B_\varphi}\right)^{5/2}\frac{m^{5/2}}{n^{1/2}} = 0, \tag{31}$$

where $b_s \equiv 2m\Psi_V/r_s$. Equation (31) differs from the cubic equation for Ω obtained by Fitzpatrick.[23] It has one physical solution, given by

$$\Omega = \frac{\Omega_0}{2}\left[1 + \sqrt{1 - \left(\frac{b_s}{b_{s,\mathrm{lock}}}\right)^{5/2}}\right], \tag{32}$$

with the locking threshold,

$$\frac{b_{s,\mathrm{lock}}}{B_\varphi} = \frac{n^{1/5}}{m}\left(\frac{\Omega_0^2\tau_A^2\tau_{SPV}}{2\lambda_{p,e}\tau_V}\right)^{2/5}\left(\frac{r_s}{R_0}\right)^{1/5}. \tag{33}$$

Note that since $\lambda_e \gg \lambda_p$, Eq. (33) yields two thresholds, one much higher than the other. It is easy to see that

$$\frac{b_{s,\mathrm{lock}}^{upper}}{B_\varphi} \sim \left(\frac{r_s}{w}\right)^{2/5}\frac{b_{s,\mathrm{lock}}^{lower}}{B_\varphi}.$$

It turns out that the upper threshold cannot be met for most parameters of physical relevance.

In the following section, we consider the different possible scenarios for the time evolution of the system.

D. Saturation levels and mode locking

1. The lower saturation level

In the presence of viscosity, the reconnection time scale τ_{SPV} in Eq. (25) is much shorter than the resistive diffusion time scale τ_R. As in the Sweet–Parker model, the reconnection dynamics is helicity conserving and is characterized by the persistence of Y points[14] until the dynamics enters a constant-ψ regime. The condition for the breakdown of the constant-ψ approximation is $\lambda\Delta' \gg 1$ or equivalently, $w\Delta' \sim 1$, where $\lambda \sim O(r_s)$ is the wavelength and w is the island width. This can be seen easily by writing the following expansion for the perturbed flux inside the island:

$$\psi(x) = \psi(0) + \psi'(0)x + O(x^2) \approx \psi(0)[1 + \Delta' x], \tag{34}$$

where $0 \leq x \leq w$. Clearly, if $w\Delta' \ll 1$, we can ignore the variation of ψ inside the island and assume $\psi(x) \approx \psi(0)$. This is the so-called constant-ψ approximation. If $w\Delta' \sim 1$ the constant-ψ approximation breaks down and the reconnection dynamics occurs in a nonconstant-ψ regime.

From the outer region solution, we obtain

$$r_s\Delta' \geq \frac{2m\Psi_V}{|\Psi_R|} = \left(\frac{B_\theta(r_s)}{b_s}\right)^{1/2} \Omega\tau_{SPV}. \tag{35}$$

Therefore, if

$$b_s < b_{s,\text{lock}}^{\text{lower}} \ll B_\theta(r_s)[\Omega\tau_{SPV}]^2,$$

the island will remain in a nonconstant-ψ regime.

Making use of Eqs. (9)–(11), we obtain

$$\frac{1}{2}s(r_s)B'_\theta(r_s)w^2 \approx 2\Psi_R + \left(\frac{2m}{r_s}\Psi_V + \psi'_m(r_s)\Psi_R\right)w$$

$$\approx 2\Psi_R + \frac{2m}{r_s}\Psi_V w. \tag{36}$$

Defining $\tilde{b}_s \equiv b_s/B_s$, and $B_s \equiv r_s s(r_s)B'_\theta(r_s)$, we can show that the island width will saturate at a low level, given by

$$\frac{w_{s,\text{lower}}}{r_s} \approx \tilde{b}_s\left[1+\left(1+\frac{4(B_s/B_\theta(r_s))^{1/2}}{\tilde{b}_s^{1/2}\Omega\tau_{SPV}}\right)^{1/2}\right]. \tag{37a}$$

We now consider special cases of Eq. (37a). If $\tilde{b}_s^{1/2}\Omega\tau_{SPV} \gg 1$, we obtain

$$\frac{w_{s,\text{lower}}}{r_s} \approx 2\tilde{b}_s, \tag{37b}$$

which yields

$$w\Delta' \approx 2\tilde{b}_s^{1/2}\left(\frac{B_\theta(r_s)}{B_s}\right)^{1/2}\Omega\tau_{SPV} \sim \tilde{b}_s^{1/2}\Omega\tau_{SPV} \gg 1.$$

If $\tilde{b}_s^{1/2}\Omega\tau_{SPV} \sim 1$, we obtain

$$\frac{w_{s,\text{lower}}}{r_s} > 2\tilde{b}_s, \tag{37c}$$

which yields

$$w\Delta' \sim \tilde{b}_s^{1/2}\Omega\tau_{SPV} \sim 1.$$

However, if $\tilde{b}_s^{1/2}\Omega\tau_{SPV} \ll 1$, we obtain

$$\frac{w_{s,\text{lower}}}{r_s} \approx 2\tilde{b}_s^{1/2}/(\Omega\tau_{SPV})^{1/2}, \tag{37d}$$

which yields

$$w\Delta' \sim (\Omega\tau_{SPV})^{1/2} \sim (\Omega\tau_A)^{1/2}S^{3/8}/R^{1/8}.$$

Simple estimates show that $w\Delta' \sim 1$ for COMPASS-C as well as ITER (see the numbers given in Sec. IV). Hence, the constant-ψ approximation is violated for these devices. For DIII-D, the error fields are ramped to attain their maximum value slowly (i.e., on a time scale comparable to the resistive diffusion time scale), invalidating the assumption of a sudden external perturbation. We postpone the correct treatment for DIII-D to a later publication.

2. Mode locking and the higher saturation level

If $b_s > b_{s,\text{lock}}^{\text{upper}}$, mode locking occurs early in the nonconstant-ψ nonlinear phase, and the reconnected flux will grow linearly with time:[12]

$$\Psi_R \approx \frac{[2m\Psi_V]^{3/2}t}{[r_sB_\theta(r_s)]^{1/2}\tau_{SPV}}. \tag{38}$$

Eventually, the island will reach a constant-ψ phase.

If $b_{s,\text{lock}}^{\text{upper}} > b_s > b_{s,\text{lock}}^{\text{lower}}$, then locking will not occur in the early nonlinear stage. Eventually, the rotation profile will relax fully and the lower locking threshold will then become the relevant one. Hence, mode locking will occur after complete relaxation. Note that since the island does not grow until it locks, it will remain in the nonconstant-ψ phase as long as it does not lock. After locking, the island will grow in accordance with Eq. (38) and evolve into a constant-ψ phase.

In the constant-ψ nonlinear phase, the island will saturate at a higher level than given by Eq. (37a). This higher level can be calculated following HK. From Ohm's law, we obtain

$$\frac{d\Psi_R}{dt} \approx \frac{r_s^2}{\tau_R}\nabla^2\psi \approx \frac{r_s^2\,\Delta b_s}{w\tau_R}, \tag{39}$$

where $w = 2[\Psi_R/s(r_s)B'_\theta(r_s)]^{1/2}$ for constant-ψ islands. Before we discuss the case of forced reconnection, we verify that the well-known Rutherford solution[15] for free reconnection caused by a constant-ψ instability follows simply from Eq. (39). Since $\Delta b_s = \Delta'_0\Psi_R$ for the instability, we obtain

$$\frac{d}{dt}\left(\frac{w}{r_s}\right) \approx \frac{\Delta'_0 r_s}{2\tau_R} = \text{const}, \tag{40}$$

which yields

$$\frac{w}{r_s} \cong \frac{\Delta'_0 r_s}{2\tau_R}t. \tag{41}$$

Now, in the case of forced reconnection, when the constant-ψ regime is accessible, we obtain

$$r_s\,\Delta b_s = 2m\Psi_V + \Psi_R r_s\psi'_m(r_s). \tag{42}$$

The island equation is then

$$\frac{dW}{dT} = \frac{1}{W^2} - 1, \tag{43}$$

where

$$W \equiv \left(\frac{r_s^3\psi'_m(r_s)s(r_s)B'_\theta(r_s)}{2m\Psi_V}\right)^{1/2}\frac{w}{2r_s}$$

and

$$T \equiv \left(\frac{r_s\psi'_m(r_s)}{2}\right)\left(\frac{r_s^3\psi'_m(r_s)B'_\theta(r_s)}{2m\Psi_V}\right)^{1/2}\frac{t}{\tau_R}.$$

According to Eq. (43), the island saturates when $W = 1$. This saturation level is generally higher than given by Eq. (37a), and given by

$$\frac{w_{s,\text{upper}}}{2r_s} = \left(\frac{\tilde{b}_s}{r_s\psi'_m(r_s)}\right)^{1/2}. \tag{44}$$

Apart from the cases discussed above [leading to saturated widths (37a) and (44)], the only remaining possibility is that in which the field error lies in the range

$$b_s > B_\theta(r_s)[\Omega\tau_{SPV}]^2. \tag{45}$$

In this case, the island may evolve into a constant-ψ regime. However, it is readily seen that the constant-ψ condition (45) requires a field error that is much too large (of the order of the toroidal field for COMPASS-C parameters!) to be relevant.

It is thus clear that for most cases of physical interest, mode locking occurs if $b_s > b_{s,\mathrm{lock}}^{\mathrm{lower}}$. For clarity, we display this locking threshold, identifying $b_{s,\mathrm{lock}} = b_{s,\mathrm{lock}}^{\mathrm{lower}}$ and replacing $\lambda_{p,e}$ by λ_e in Eq. (33):

$$\frac{b_{s,\mathrm{lock}}}{B_\varphi} = \frac{n^{1/5}}{m}\left(\frac{\Omega_0^2 \tau_A^2 \tau_{SPV}}{2\lambda_e \tau_V}\right)^{2/5}\left(\frac{r_s}{R_0}\right)^{1/5}. \tag{46}$$

We note that the parametric dependencies of Eq. (46) are different from those given by the constant-ψ theory. For instance, the threshold (46) scales with the (electron) density as $n_e^{1/2}$, which differs from the $n_e^{2/3}$ scaling obtained by FH.[4] (This assumes a deuterium plasma with equal average electron and ion densities and ignores the possible density dependencies of τ_R and τ_V.)

E. Summary

The results of Sec. III can be summarized as follows.

The linear regime is nonconstant-ψ, with a time scale $\tau_{LD} = (\tau_A^2 \tau_R^2/\tau_V)^{1/3} \ll \Omega_0^{-1}$. The island rotation frequency slows down, but does not lock to the error field. The transition to the nonlinear stage occurs in a nonconstant-ψ regime, characterized by a characteristic reconnection time scale τ_{SPV}, helicity conservation, and the persistence of current sheets. In the nonconstant-ψ regime, the locking threshold is given by Eq. (46). If

(1) $b_s < b_{s,\mathrm{lock}} \ll B_\theta(r_s)[\Omega \tau_{SPV}]^2$, the island will saturate at the level (37a).

(2) $b_s > b_{s,\mathrm{lock}}$, mode locking will occur, and the island will evolve into a constant-ψ phase. In the constant-ψ phase, the island will saturate at the level (44).

IV. IMPLICATIONS FOR EXPERIMENTS

We now consider the implications of the theoretical results obtained above for experimental results from COMPASS-C. For devices such as DIII-D in which the error fields are ramped to attain their maximum value slowly (i.e., on a time scale comparable to the resistive diffusion time scale), the discussion in Secs. II and III is not applicable. In COMPASS-C, error fields are rapidly enhanced by the deliberate imposition of resonant magnetic field perturbations via external field coils. (See, for instance, Fig. 2 of Ref. 5.) Hence, the results of Sec. III can be applied to COMPASS-C.

In Fig. 1, we plot the experimental data (taken from Ref. 4) on the critical density for locked modes in COMPASS-C as a function of the measured $m=2$, $n=1$ error field. The solid line indicates the theoretical predictions fitted to the formula (46). In making these fits we have taken[4,5] $\Omega_0 = 10^5$/s, $\tau_A = 10^{-7}$ s, $S = 10^6$, and $R = 10^5$. (It should be borne in mind that the theoretical formulas have been derived in the regime $S \gg R$.) For these parameters, with $n_e \sim 10^{13}$ cm^{-3}, we use (46) to obtain the threshold

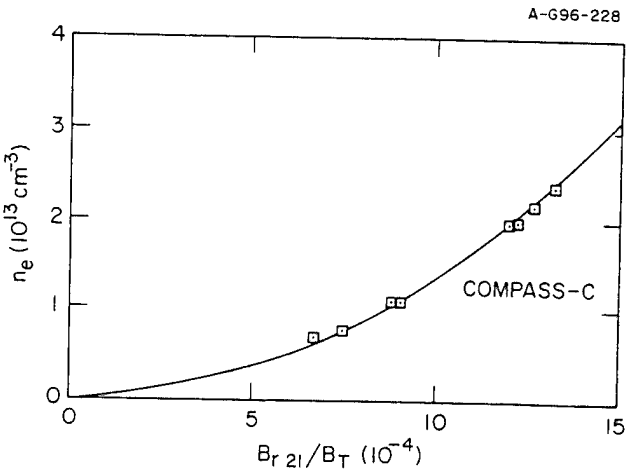

FIG. 1. Observed critical density of locked modes for Ohmic, deuterium discharges in COMPASS-C as a function of the $m=2$, $n=1$ radial error field (normalized by the toroidal field at the edge). The solid lines are the theoretical predictions. The theory curves are computed using Eq. (46). For other plasma parameters, see Sec. IV.

$b_{s,\mathrm{lock}}/B_\varphi \approx 3 \times 10^{-3}$. Since there are uncertainties in some of these assumed parameters, the fit is done by assuming that the first point of the experimental curve coincides with the relevant theoretical formula (which is scaled by a multiplicative constant of order unity); the theoretical formula is then used to predict the remaining points.

The establishment of a steady-state plasma equilibrium in the presence of a pre-existing error field has not been considered in this paper. (In its full complexity, the problem is fully three dimensional, as in stellarators.[26]) Strictly speaking, one cannot use the results of this paper to make reliable predictions for critical pre-existing error fields in the ITER device. However, if we follow Ref. 4 (where the FH theory is used) and do so, the results of Sec. III can be used to make comparisons with the predictions of FH. The critical error field for locked modes in ITER predicted by FH is quite stringent, given by $B_{r21}/B_T \approx 2 \times 10^{-5}$.[4] Following Ref. 4, we take $\Omega_0 \tau_A \sim 10^{-4}$ and $r_s/R_0 \approx 0.3$ for ITER. We assume $S \sim 10^9$ and $R \sim 10^8$, consistent with the theoretical assumption $R \ll S$. For these numbers, we obtain the threshold $B_{r21}/B_T \approx 6 \times 10^{-5}$, which is a little less stringent than the prediction of FH.

V. SUMMARY

An important mechanism for mode-locking in cylindrical plasmas is forced reconnection due to resonant magnetic field perturbations. The dynamics of this process in the linear and nonlinear regimes are generally characterized by very large values of the parameter Δ' that invalidate the constant-ψ approximation, used in the standard model of Fitzpatrick and Hender.[16] We have revisited this model and obtained results in the nonconstant-ψ linear and nonlinear regimes. As shown in Ref. 19 as well as this paper, the results of the nonconstant-ψ theory are qualitatively different, and are

characterized by the features of fast reconnection and persistence of current sheets (spanning Y points). Such structures cannot be dealt with by a constant-ψ theory.

The theory presented in this paper gives a fairly complete picture of the time history of forced reconnection in the linear as well as nonlinear regimes, driven by an error field that is switched on suddenly in rotating cylindrical plasmas. We find that the time history as well as the critical thresholds are different from those obtained from a constant-ψ theory. For critical thresholds, the principal result is given by Eq. (46), which has been used for comparison with data from COMPASS-C. The application of the present theory to the observations in DIII-D is complicated by the slow ramp of the error field and the possibility of formation of constant-ψ islands in the early nonlinear stage. This will be the subject of a separate publication.

ACKNOWLEDGMENTS

We thank Professor R. Fitzpatrick for useful dicussions.

This research is supported by the Air Force Office of Scientific Research Grant No. F49620-93-1-10071 and the National Science Foundation Grants No. ATM 93-10157 and No. ATM 93-96280.

APPENDIX: ON THE BREAKDOWN OF THE CONSTANT-ψ APPROXIMATION IN THE LINEAR REGIME OF FORCED RECONNECTION

It has been clearly established that in the early linear phase, forced reconnection will begin in a nonconstant-ψ stage with an initial condition of zero reconnected flux in the resistive (HK and WB) as well as viscoresistive limits. [See, for example, Fig. 2 and Eq. (25) of HK, Fig. 4 of FH, obtained by numerical integration of the linearized equations, and Eqs. (16), (40), and (54) of ZWB, who also present numerical tests of their analytical results.] However, in Sec. 7 of Ref. 17, Fitzpatrick claims that "the constant-ψ approximation holds in the viscoresistive limit" of forced reconnection because of the condition [inequality (115) of Ref. 17]

$$|\Delta\omega|\tau_R\left(\frac{\delta_{\text{layer}}}{r_s}\right)^2 \ll 1. \tag{A1}$$

Comparing Eqs. (111a) and (111b) of Ref. 17 to Eqs. (20)–(22) of the present paper, we see readily that they are identical, except that the variable s (or equivalently, $\partial/\partial t$) in the present paper is replaced in by $i\,\Delta\omega$ in Ref. 17. Note that $\Delta\omega$ is equivalent to our parameter Ω_0. Then, the essential difference between our treatment and that of Ref. 17 can be traced to the assumption in Ref. 17 that the functions $\hat{\psi}$ and $\hat{\phi}$ (in our notation) are independent of time. Such an assumption is typical for problems involving free reconnection or tearing eigenmodes, and fundamentally incorrect for forced reconnection. During free reconnection, an unstable eigenmode can grow exponentially in time from an arbitrarily small nonzero perturbation of the equilibrium. During forced reconnection, the perturbation grows out of the stable MHD continuum, and one has to solve an initial-value problem with an initial condition of zero reconnected flux. Therefore, the term $i\,\Delta\omega$ in Sec. 7 of Ref. 17 should be replaced by the variable $s=p-i\,\Delta\omega$. Then the dimensionless quantity on the left-hand side of Eq. (A1) [or inequality (115) of Ref. 17] can be rewritten as

$$s\tau_R\epsilon_D^2 \equiv s\epsilon_D^2/\epsilon^3 \equiv \mu_D = s\tau_{LD}, \tag{A2}$$

where $\tau_{LD}=(\tau_A\tau_R)^{2/3}/\tau_V^{1/3}$ (see Sec. IV of ZWB). It follows then there should always exist a period $t<\tau_{LD}$, or $\mu_D>1$ when the constant-ψ approximation (A1) breaks down. In fact, as shown by ZWB, the linear phase in the viscoresistive regime holds when $t\ll\tau_{LD}$, or $\mu_D\gg 1$, and that the reconnection dynamics enters a nonlinear nonconstant-ψ regime, with no room whatsoever for a constant-ψ linear phase when $\mu_D\ll 1$.

[1] J. A. Snipes, D. J. Campbell, T. C. Hender, M. von Hellermann, and H. Weisen, Nucl. Fusion **30**, 205 (1990).
[2] J. Chen, J. Xie, Y. Huo, Lin. Li, Q. Zhao, G. Zhang, M. Wang, D. Guo, Q. Guo, P. Qin, G. Li, H. Fan, C. Deng, X. Tong, and R. Huang, Nucl. Fusion **30**, 2271 (1990).
[3] J. T. Scoville, R. J. La Haye, A. G. Kellman, T. H. Osborne, R. D. Stambaugh, E. J. Strait, and T. S. Taylor, Nucl. Fusion **31**, 875 (1991).
[4] R. J. La Haye, R. Fitzpatrick, T. C. Hender, A. W. Morris, J. T. Scoville, and T. N. Todd, Phys. Fluids B **4**, 2098 (1992).
[5] T. C. Hender, R. Fitzpatrick, A. W. Morris, P. G. Carolan, R. D. Durst, T. Edlington, J. Ferreira, S. J. Fielding, P. S. Haynes, J. Hugill, I. J. Jenkins, R. J. La Haye, B. J. Parham, D. C. Robinson, T. N. Todd, M. Valovic, and G. Vayakis, Nucl. Fusion **32**, 2091 (1992).
[6] R. J. La Haye, A. W. Hyatt, and J. T. Scoville, Nucl. Fusion **32**, 2119 (1992).
[7] A. F. Almagri, S. Assadi, S. C. Prager, J. S. Sarff, and D. W. Kerst, Phys. Fluids B **4**, 4080 (1992).
[8] G. M. Fishpool and P. S. Haynes, Nucl. Fusion **34**, 109 (1994).
[9] D. J. Den Hartog, A. F. Almagri, J. T. Chapman, H. Ji, S. C. Prager, J. S. Sarff, R. J. Fonck, and C. C. Hegna, Phys. Plasmas **2**, 2281 (1995).
[10] T. H. Ivers, E. Eisner, A. Garofalo, R. Kombargi, M. E. Mauel, D. Maurer, D. Nadle, G. A. Navratil, M. K. V. Sankar, M. Su, E. Taylor, Q. Xiao, R. R. Bartsch, W. A. Reass, and G. A. Wurden, Phys. Plasmas **3**, 1926 (1996).
[11] T. S. Hahm and R. M. Kulsrud, Phys. Fluids **28**, 2412 (1985).
[12] X. Wang and A. Bhattacharjee, Phys. Fluids B **4**, 1795 (1992).
[13] W. Park, D. A. Monticello, and R. B. White, Phys. Fluids **27**, 137 (1984).
[14] F. L. Waelbroeck, Phys. Fluids B **1**, 2372 (1989).
[15] P. H. Rutherford, Phys. Fluids **16**, 1903 (1973).
[16] R. Fitzpatrick and T. C. Hender, Phys. Fluids B **3**, 644 (1991).
[17] R. Fitzpatrick, Nucl. Fusion **33**, 1049 (1993).
[18] R. Fitzpatrick, R. J. Hastie, T. J. Martin, and C. M. Roach, Nucl. Fusion **33**, 1533 (1993).
[19] Z. W. Ma, X. Wang, and A. Bhattacharjee, Phys. Plasmas **3**, 2427 (1996).
[20] J. P. Christiansen, J. G. Cordey, and K. Thomassen, Nucl. Fusion **30**, 1183 (1990).
[21] A. W. Morris, R. Fitzpatrick, P. S. Haynes, T. C. Hender, J. Hugill, C. Silvester, and T. N. Todd, in *Controlled Fusion and Plasma Heating 1990*. Proceedings of the 17th European Conference, Amsterdam (European Physical Society, Petit-Lancy, 1990), Vol. 1, p. 379.
[22] R. Fitzpatrick, in *Theory of Fusion Plasmas*, edited by E. Sindoni and J. Vaclavik (Editrice Compositori, Bologna, 1992), p. 147.
[23] R. Fitzpatrick ("Driven reconnection in magnetic fusion experiments," Lecture given at Summer School on "MHD phenomena in plasmas," Madison, W, 14–18 August 1995; private communication).
[24] P. A. Sweet, *Electromagnetic Phenomena in Cosmical Physics* (Cambridge University Press, New York, 1958).
[25] E. N. Parker, J. Geophys. Res. **62**, 509 (1957).
[26] See, for instance, A. Bhattacharjee, T. Hayashi, C. C. Hegna, N. Nakajima, and T. Sato, Phys. Plasmas **2**, 883 (1995), and other references therein.

THE ASTROPHYSICAL JOURNAL, 420:415–421, 1994 January 1
© 1994. The American Astronomical Society. All rights reserved. Printed in U.S.A.

CURRENT SHEETS AND RECONNECTION DRIVEN BY FOOTPOINT MOTION IN TWO-DIMENSIONAL CORONAL LOOPS WITH X-TYPE NEUTRAL LINES

XIAOGANG WANG AND A. BHATTACHARJEE
Department of Applied Physics, 500 West 120th Street, Columbia University, New York, NY 10027
Received 1993 April 26; accepted 1993 July 12

ABSTRACT

The dynamical evolution of a two-dimensional coronal loop from a smooth initial state containing an X-type neutral line is considered. An exact solution of the linearized ideal magnetohydrodynamic equations shows that the amplitude of the current sheet at the separatrix grows exponentially with time while its width shrinks at the same rate. Thus, although there is a strong tendency for a current sheet to form, a true singularity is not realized in finite time. Resistivity intervenes, and the ideal phase is followed by a linear resistive phase in which the dynamics is still exponential in time with a growth rate proportional to $S^{-1/3}$, where S is the Lundquist number of the coronal plasma. The linear resistive phase is followed by a helicity-conserving nonlinear phase in which the growth rate is algebraic in time, and the reconnection rate is proportional to $S^{-1/2}$, as in the Sweet-Parker model. It is demonstrated that the heating caused by these current sheets can be large enough to account for the energy balance in quiet as well as some active coronal loops.

Subject headings: MHD — Sun: corona — Sun: magnetic fields — Sun: X-rays, gamma rays

1. INTRODUCTION

For almost two decades now, the problem of current sheet formation in the solar corona has stimulated considerable research as well as controversy. There are two facets to this problem: one involves geometry, and the other involves dynamics. Geometry is involved in the question: *where* do current sheets form? Dynamics determines the answer to the question: *how rapidly* in time do current sheets form?

There are some magnetic configurations in which the question pertaining to geometry has a clear answer. For instance, it is well-understood that the natural sites for singularity formation in a toroidal configuration are the so-called rational surfaces on which field lines close on themselves. In this case, the rational surfaces are the source of separatrices. Dynamical considerations, based upon ideal and resistive magnetohydrodynamics (MHD), show that current singularities tend to form at the separatrices, but not in finite time (Rosenbluth, Dagazian, & Rutherford 1973; Hahm & Kulsrud 1985; Waelbroeck 1989; Wang & Bhattacharjee 1992a). [If there is a finite-time singularity, the current density may blow up, for instance, as $(t_c - t)^{-\alpha}$, where t_c and α are positive constants. In contrast, if the current grows algebraically or exponentially with time, it is an infinite-time singularity.]

If we consider three-dimensional magnetic fields in general, it is not obvious where the separatrices are. Greene (1988) has emphasized that magnetic nulls (where the magnetic field vanishes identically) are a copious source of separatrices. But there are many configurations of relevance to the solar corona in which the magnetic field may not vanish anywhere. A popular model, suggested by Parker (1972), involves a straight axial magnetic field $\boldsymbol{B} = B_0 \hat{z}$ between two parallel plates at $z = \pm L$ which is then deformed by the motion of photospheric footpoints in the x-y plane. If the length L of a straight, cylindrical coronal tube is much larger than its transverse dimension a, it is possible to calculate some asymptotic force-free equilibria using the inverse aspect ratio a/L as a small parameter (van Ballegooijen 1985; Zweibel & Li 1987; Strauss & Otani 1988; Bhattacharjee & Wang 1991; Wang & Bhattacharjee 1992b; Longcope & Sudan 1992). (This "long-thin" approximation is appropriate for many coronal loops for which typically $L \sim 10^{10}$ cm, and $a \sim 10^8$ cm which yields an aspect ratio $L/a \sim 100$.) Then, if the z-dependence of $\boldsymbol{B}$ is weak, the effective dimensionality of the problem can be shown to reduce from three to two. This reduction leads to considerable simplification of the three-dimensional MHD equations. Elsewhere, we have shown that free or forced reconnection processes which conserve magnetic helicity tend to produce current sheets near the separatrices of a long-thin coronal tube (Bhattacharjee & Wang 1991; Wang & Bhattacharjee 1992b).

Although the long-thin coronal tube is a useful model, it entails an asymptotic reduction of the three-dimensional MHD equations keeping B_z large and fixed. The relevance of the conclusions obtained from this reduced MHD model to the solar corona can be questioned on the ground that the footpoint motions on the photosphere are likely to affect B_z. Also, these are asymptotic magnetostatic equilibria, and the structure of separatrices in such a model depends sensitively on the approximations used.

A different model of the solar corona that is relevant and yet has the virtue of analytical tractability is shown in Figure 1. (The x and y coordinates lie in the plane of the paper and the z-coordinate points upward.) The system is invariant to translation along the z-axis and is characterized by the presence of an X-type neutral line. The formation of current sheets in this geometry has been investigated analytically as well as numerically (Zwingmann, Schlindler, & Birn 1985; Antiochos 1987; Moffatt 1987; Aly 1987; Low & Wolfson 1988; Jensen 1989; Karpen, Antiochos, & DeVore 1990; Zweibel & Proctor 1990; Finn & Lau 1991; Vekstein & Priest 1992). It turns out that this simple (but subtle!) model is not free of controversy. Whereas some (Zwingmann et al. 1985; Moffatt 1987; Aly 1987; Low & Wolfson 1988; Jensen 1989; Zweibel & Proctor 1990) claim that current sheets should occur at the separatrix by continuous and finite footpoint motions in an initially smooth equilibrium, others (Karpen et al. 1990; Finn & Lau 1991) refute such a general claim. We review here briefly the different aspects of this debate.

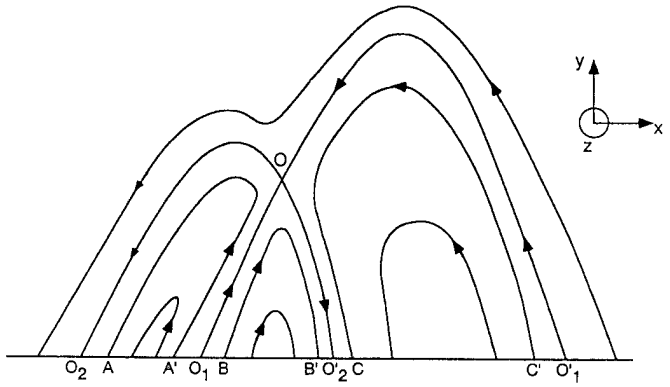

FIG. 1.—Projection of constant-ψ surfaces on the x-y plane for a coronal loop with an X-type neutral line.

For a configuration with translation symmetry in z, the magnetic field $\boldsymbol{B}$ can be represented as

$$\boldsymbol{B} = \nabla\psi(x, y) \times \hat{z} + B_z(x, y)\hat{z} . \tag{1}$$

The Grad-Shafranov equation for static force-free equilibria is

$$\nabla^2\psi + \frac{d}{d\psi}\frac{B_z^2(\psi)}{2} = 0 , \tag{2}$$

where $B_z = B_z(\psi)$. In Figure 1, we show a cross section of a coronal loop in the x-y plane, with the footpoints of the field lines intersecting the photosphere which is taken to be the $y = 0$ plane. The separatrices $O_1O'_1$ and $O_2O'_2$ cross at O which is the point of intersection of the X-type neutral line with the x-y plane.

Let us assume that there is a smooth equilibrium with $B_z = 0$ everywhere initially. Then the motion of the footpoints in the z-direction will cause the growth of B_z. If we assume that the motion of footpoints takes the configuration through a sequence of static equilibria, then these equilibria can be calculated by specifying the footpoint displacement in the z-direction,

$$d(\psi) = B_z(\psi)v'(\psi) , \tag{3}$$

where $v'(\psi) = \int dl/B$. An important question on which opinion is divided is the following: is it possible for this configuration to develop current sheets at the separatrix by means of *smooth* photospheric displacements?

Some answer the question in the affirmative. They note that field lines to the left of O_1, such as the one at A', are connected to points to the right of O_2, such as A. This means that $B_z(A) = B_z(A')$, and $d(A) = d(A')$. On the other hand, field lines to the right of O_1, such as the one at B, are connected to points to the left of O'_2. This implies that $B_z(B) = B_z(B')$ and $d(B) = d(B')$. However, since $d(A)$ and $d(B)$ are independent of each other, it is possible for them to be discontinuous across the separatrix for smooth motions of the footpoints. For example, let us assume that the footpoints near O'_2 move in the z-direction to yield $B_z(B) = B_z(B') > 0$ and $B_z(C) = B_z(C') > 0$, but that the footpoints to the left of O_1 and to the right of O_2 do not move which means that $B_z(A) = B_z(A') = 0$. Then as A' and $B \to O_1$ and $B' \to O_2$, there should be a jump in B_z and a current sheet on the separatrix, given by

$$\Delta B_z = B_z(B) - B_z(A') = B_z(B) > 0 . \tag{4}$$

The simple argument given above is complicated by the fact that in this class of force-free equilibria with $B_z = B_z(\psi)$, there may be logarithmic singularities in the photospheric displacements $d(\psi)$. Assuming that the point O is hyperbolic and that $\psi = 0$ at O, it can be shown that $v'(\psi) \sim \ln|\psi|$ (Grad, Hu, & Stevens 1975; Zwingmann et al. 1985). If we take $B_z \simeq B_0 =$ constant, it follows that $d(\psi) \sim \ln|\psi|$ which means that $d(\psi)$ must have logarithmic singularities as the separatrix $\psi = 0$ is approached. [If, on the other hand, we demand that $d(\psi) \to d_0 =$ constant as $\psi \to 0$, we get $d/d\psi(B_z^2/2) \sim 1/(\psi \ln|\psi|)^3$ which indicates that it would be impossible to satisfy the condition for force-free equilibria.]

To avoid the logarithmic singularities in $d(\psi)$, Finn & Lau (1991) construct classes of self-similar equilibrium solutions of the form $\psi = r^{2\beta}Q(\theta)$, where $r^2 = x^2 + y^2$, the angle θ is measured from one arm of the separatrix where the function Q vanishes and β is a constant that lies in the range $\frac{1}{2} < \beta < 1$. These solutions have the virtue that the footpoint displacement $d(\psi)$ is a continuous function in each of the quadrants divided by the separatrix. Moreover, the current density J_z is not a δ-function but has fractional power-law singularities at the separatrix.

Though interesting in their own right, Finn & Lau's solutions (and generalizations thereof, due to Vekstein & Priest 1992) are derived entirely from considerations of magnetostatics. The assumption is made implicitly that the coronal plasma will adjust to photospheric footpoint motion in order to maintain magnetostatic equilibrium, but no attempt is made to determine how rapidly in time the adjustment actually occurs. One of the main objectives of this paper is to examine critically this assumption, for it is at the heart of the question of *dynamical accessibility* of any singular equilibrium from a smooth initial state.

The solutions of Finn & Lau depend crucially on the validity of the condition $B_z = B_z(\psi)$ which is a requirement for magnetostatic equilibrium. Let us consider a smooth initial equilibrium for which $B_z = 0$. If this equilibrium is deformed by smooth footpoint displacements, it is important to determine from the dynamical equations of ideal MHD whether the condition $B_z = B_z(\psi)$ is attained in finite time. One of the significant results of this paper is that the state $B_z = B_z(\psi)$ is not attainable in finite time. An exact solution of the (linearized) ideal MHD equations shows that for short times, of the order of the Alfvén timescale, the amplitude of the current sheet grown exponentially with time while its width shrinks exponentially (§ 3). Thus a current sheet of truly infinite amplitude and zero thickness is not dynamically accessible in finite time if we begin from a smooth initial state.

What, then, is the role of Finn & Lau's equilibrium solutions with power-law singularities in the actual dynamical evolution of the corona? If we choose as our initial condition one of these equilibrium solutions, we show (in the Appendix) that current singularities are realized in finite time. But this occurs because the singularity already exists in the initial state! If we demand, as is physically reasonable, that the initial state be smooth, then none of the solutions with power-law singularities are dynamically accessible in finite time.

Resistivity, no matter how small, intervenes in finite time. Beginning from a smooth initial state, we demonstrate that the ideal phase in which the current sheet amplitude grows exponentially is followed by a linear resistive phase in which the amplitude continues to grow exponentially, but proportional to $\eta^{1/3}$ where η is the plasma resistivity (§ 4). This linear phase is then followed by a helicity-conserving nonlinear phase in

which the growth of the current sheet is algebraic in time, and the reconnection rate is proportional to $\eta^{1/2}$, as in the Sweet-Parker model (Sweet 1958; Parker 1957) (§ 5). In this phase, due to the constraint imposed by helicity conservation (Waelbroeck 1989; Bhattacharjee & Wang 1991), the X-point structure in the initial state changes to a Y-point structure (Sweet 1958; Syrovatsky 1981). Estimates of coronal heating due to current sheets along X-type neutral lines shows that the heating is large enough to account for the energy balance in quiet, as well as some active coronal loops. In § 6, we give a summary of our results and discuss their implications.

2. THE DYNAMICAL EQUATIONS

For footpoint displacements which preserve the translation invariance of the system along z, the magnetic field $\boldsymbol{B}$ may be calculated using the representation (1). The time evolution of ψ and B_z are specified by the equations

$$\frac{d\psi}{dt} = \frac{\eta c^2}{4\pi} \nabla^2 \psi \,, \tag{4}$$

and

$$\frac{dB_z}{dt} = (\psi, v_z) + \frac{\eta c^2}{4\pi} \nabla^2 B_z \,, \tag{5}$$

where $d/dt = \partial/\partial t + \boldsymbol{v}_\perp \cdot \nabla$, and

$$(f, g) = \hat{z} \cdot \nabla f \times \nabla g = \frac{\partial f}{\partial x}\frac{\partial g}{\partial y} - \frac{\partial f}{\partial y}\frac{\partial g}{\partial x} \,. \tag{6}$$

The fluid velocity v can be represented as

$$\boldsymbol{v} = \hat{z} \times \nabla \phi + v_z \hat{z} \,. \tag{7}$$

From the momentum equation, we get

$$\frac{d\omega}{dt} = \frac{1}{c\rho}(\psi, J) \,, \tag{8}$$

and

$$\frac{dv_z}{dt} = \frac{1}{4\pi\rho}(\psi, B_z) \,, \tag{9}$$

where $\omega = \nabla^2 \phi$, $J = (c/4\pi)\nabla^2 \psi$, ρ is a constant density, and

$$\frac{d}{dt} = \frac{\partial}{\partial t} + (\phi, \quad) \,. \tag{10}$$

To simplify the analysis, we consider the geometry drawn in Figure 2 which is topologically equivalent to the one in Figure 1. In particular, the X-type neutral line which intersects the x-y plane at the point O, is contained in the box (of width a) indicated by the dotted line in Figure 2. The initial equilibrium is taken to be a vacuum magnetic field described by

$$\psi_0(x, y) = B_0 xy/a \,. \tag{11}$$

[Note that $J_0 = (c/4\pi)\nabla^2 \psi_0 = 0$.] The condition for ideal magnetostatic equilibrium gives

$$(\phi_0, \psi_0) = (\phi_0, \omega_0) = 0 \,, \tag{12}$$

which is obtained by setting $\partial/\partial t = 0$ and $\eta = 0$ in equations (4) and (8). Equation (12) implies that

$$\phi_0 = \phi_0(\psi_0) \,, \quad \omega_0 = \omega_0(\psi_0) \,. \tag{13}$$

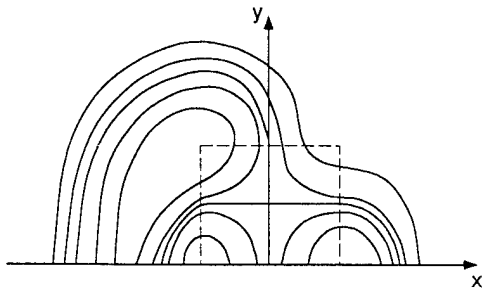

FIG. 2.—Projection of constant-ψ surfaces on the x-y plane for a model initial equilibrium. The contours within the dotted square correspond to $\psi_0 = B_0 xy/a$.

Furthermore, equations (5) and (9) give, respectively,

$$(\psi_0, B_{z0}) + [\phi_0(\psi_0), v_{z0}] = 0 \,, \tag{14}$$

and

$$(\psi_0, v_{z0}) + [\phi_0(\psi_0), B_{z0}] = 0 \,. \tag{15}$$

If we define the equilibrium Elsässer variables $f_0 = B_{z0} + v_{z0}$ and $g_0 = B_{z0} - v_{z0}$, then equations (14) and (15) can be combined to give

$$(\psi_0, f_0) + [\phi_0(\psi_0), f_0] = 0 \,, \tag{16}$$

and

$$(\psi_0, g_0) - [\phi_0(\psi_0), g_0] = 0 \,. \tag{17}$$

Equations (16) and (17) are satisfied if

$$f_0 = f_0(\psi_0) \quad \text{and} \quad g_0 = g_0(\psi_0) \,, \tag{18}$$

or if

$$B_{z0} = B_{z0}(\psi_0) \quad \text{and} \quad v_{z0} = v_{z0}(\psi_0) \,. \tag{19}$$

Equation (19) must be satisfied by all equilibrium solutions. We demonstrate in § 3 that if the initial equilibrium (11) is perturbed, then the condition (19) cannot be attained in finite time. For simplicity, we consider footpoint displacements along z that perturb dominantly the equilibrium magnetic field in the square marked by the dotted line as in Figure 2. We focus on the square because it contains the X-line.

3. THE LINEAR IDEAL PHASE: AN EXACT SOLUTION

In what follows, it is convenient to transform to dimensionless variables $\boldsymbol{B}/B_0 \to \boldsymbol{B}$, $\boldsymbol{x}/a \to \boldsymbol{x}$, $t/\tau_A \to t$, where $\tau_A = a/v_A$ and $v_A = B_0/(4\pi\rho)^{1/2}$ is the Alfvén speed. For the initial equilibrium (11), since $B_{z0} = v_{z0} = \phi_0 = 0$, the linearized ideal MHD equations for B_z and v_z are given, respectively, by

$$\frac{\partial B_z}{\partial t} = y\frac{\partial v_z}{\partial y} - x\frac{\partial v_z}{\partial x} \,, \tag{20}$$

and

$$\frac{\partial v_z}{\partial t} = y\frac{\partial B_z}{\partial y} - x\frac{\partial B_z}{\partial x} \,. \tag{21}$$

If we define the Elsässer variables $f = B_z + v_z$ and $g = B_z - v_z$, then equations (20) and (21) can be combined to give

$$\frac{\partial f}{\partial t} = y\frac{\partial f}{\partial y} - x\frac{\partial f}{\partial x} \,, \tag{22}$$

and

$$\frac{\partial g}{\partial t} = x\frac{\partial g}{\partial x} - y\frac{\partial g}{\partial y}. \tag{23}$$

Equations (22) and (23) can be integrated easily along their characteristics. For equation (22), the characteristics are

$$\frac{dx}{dt} = x, \tag{24a}$$

$$\frac{dy}{dt} = -y, \tag{24b}$$

whereas, for equation (23), they are

$$\frac{dx}{dt} = -x, \tag{25a}$$

$$\frac{dy}{dt} = y, \tag{25b}$$

Along their respective characteristics, equations (22) and (23) become

$$\frac{df}{dt} = 0, \tag{26}$$

and

$$\frac{dg}{dt} = 0. \tag{27}$$

The exact solution of equation (26) is

$$f(x, y, t) = f(x_0, y_0, t = 0) \equiv f_0(x_0, y_0), \tag{28}$$

where

$$x(t) = x_0 e^t, \tag{29a}$$

and

$$y(t) = y_0 e^{-t}, \tag{29b}$$

Equation (28) can then be rewritten as

$$f(x, y, t) = f_0(xe^{-t}, ye^t). \tag{30}$$

Similarly, from equations (25a), (25b) and (27), we get

$$g(x, y, t) = g_0(xe^t, ye^{-t}). \tag{31}$$

We define

$$b(x, y) = B_z(x, y, t = 0), \tag{32}$$

and

$$u(x, y) = v_z(x, y, t = 0), \tag{33}$$

whence

$$\begin{aligned} B_z(x, y, t) &= (f + g)/2 \\ &= \tfrac{1}{2}[b(xe^{-t}, ye^t) + u(xe^{-t}, ye^t)]. \end{aligned} \tag{34}$$

The transverse current density (i.e., the current density perpendicular to $\hat{z}$) is

$$\begin{aligned} J_\perp &= \left[\left(\frac{\partial B_z}{\partial x}\right)^2 + \left(\frac{\partial B_z}{\partial y}\right)^2\right]^{1/2}, \\ &= \left[\left(\frac{\partial B_z}{\partial x_0}\right)^2 e^{-2t} + \left(\frac{\partial B_z}{\partial y_0}\right)^2 e^{2t}\right]^{1/2}. \end{aligned} \tag{35}$$

For large t, the first term in the parentheses on the right-hand side of equation (35) is much smaller than the second term. Hence

$$J_\perp \simeq J_0 e^t, \tag{36}$$

where $J_0 \simeq \partial B_z/\partial y_0$. This current density is localized near the separatrix $\psi = 0$, and the width of the current channel shrinks as e^{-t} along y. Thus, a true current singularity is realized only in infinite time. The boundary displacement

$$d \sim \int_0^t v_z(x, y, t')dt'$$

is bounded for all time and tends to infinity as $t \to \infty$. We demonstrate in § 4 that resistivity, no matter how small, intervenes in finite time to thwart the formation of the current singularity.

Why does it take infinite time to form a current singularity in this system? In the initial state (11), $B_z = 0$ everywhere. It takes the coronal field approximately an Alfvén time $\tau_A \sim a(4\pi\rho)^{1/2}/B_\perp$ to respond to photospheric footpoint motion occurring on the timescale τ_0. If $\tau_A \ll \tau_0$, then the system adjusts to maintain magnetostatic equilibrium. However, since $B_\perp \to 0$ and $\tau_A \to \infty$ in the vicinity of the X-point, it takes the system infinite time to attain magnetostatic equilibrium.

4. THE LINEAR RESISTIVE PHASE

In the linear resistive phase, the equations governing B_z and v_z can be written (in dimensionless form) as

$$\frac{\partial B_z}{\partial t} = y\frac{\partial v_z}{\partial y} - x\frac{\partial v_z}{\partial x} + \frac{1}{S}\nabla^2 B_z, \tag{37}$$

and

$$\frac{\partial v_z}{\partial t} = y\frac{\partial B_z}{\partial y} - x\frac{\partial B_z}{\partial x}, \tag{38}$$

where $S = \tau_R/\tau_A$ is the Lundquist number, with $\tau_R \equiv 4\pi a^2/\eta c^2$. In the linear ideal phase, if $\partial/\partial x_0 \sim \partial/\partial y_0$ initially, we have

$$\frac{\partial}{\partial x} = e^{-t}\frac{\partial}{\partial x_0} \ll \frac{\partial}{\partial y} = e^t\frac{\partial}{\partial y_0}. \tag{39}$$

Hence, carrying over the approximation $\partial/\partial x \ll \partial/\partial y$ into the linear resistive phase, equations (37) and (38) simplify to

$$\frac{\partial B_z}{\partial t} \simeq y\frac{\partial v_z}{\partial y} + \frac{1}{S}\frac{\partial^2 B_z}{\partial y^2}, \tag{40}$$

and

$$\frac{\partial v_z}{\partial t} \simeq y\frac{\partial B_z}{\partial y}. \tag{41}$$

We solve the initial-value problem using Laplace transforms (Hahm & Kulsrud 1985; Wang & Bhattacharjee 1992a, b). We define

$$\tilde{B} = \int_0^\infty dt\, e^{-pt} B_z\,, \quad \mathrm{Re}\ p > 0\,. \tag{42}$$

Equations (40) and (41) yield

$$p^2(\tilde{B} - \tilde{b}) = \frac{p}{S}\frac{\partial^2 \tilde{B}}{\partial y^2} + y\frac{\partial}{\partial y}\left(y\frac{\partial \tilde{B}}{\partial y} + u\right), \tag{43}$$

where $\tilde{b} \equiv b/p$, and b and u are defined by equations (32) and (33), respectively. Following Hahm & Kulsrud (1985), we define the small parameter $\epsilon \equiv (p/S)^{1/4}$ and the stretched variable $\theta \equiv y/\epsilon$. Then, equation (43) becomes

$$p^2(\tilde{B} - \tilde{b}) = \epsilon^2 \tilde{B}'' + \theta(\tilde{B}' + u') + \theta^2 \tilde{B}''\,, \tag{44}$$

where prime denotes differentiation with respect to θ. Equation (44) can be solved in different asymptotic limits. For times $t \ll \tau_1 = S^{1/3}$ (which corresponds to $\tau_1 = \tau_R^{1/3}\tau_A^{2/3}$ in dimensional time), since $pS^{1/3} \gg 1$ for fixed S, equation (44) becomes

$$(pS^{1/3})^{3/2}(\tilde{B} - \tilde{b}) = \tilde{B}''\,. \tag{45}$$

Equation (45) motivates the choice of rescaled variables $p_1 = pS^{1/3}$ and $\theta_1 = yS^{1/3}$. Then (43) can be written in the form

$$p_1(\tilde{B}_1 - \tilde{b}_1) = \partial^2 \tilde{B}_1/\partial\theta_1^2\,, \tag{46}$$

where

$$\tilde{B}_1 = \int_0^\infty dt_1\, e^{-p_1 t_1} B_z\,, \quad t_1 \equiv t/\tau_1\,, \tag{47}$$

and $\tilde{b}_1 = b/p_1$. Writing $\tilde{B}_1 = \tilde{B}_k \exp(-k\theta_1)$, with $k > 0$, we get

$$\tilde{B}_k \simeq \frac{b}{p_1 - k^2}\,. \tag{48}$$

The inverse Laplace transform of equation (48) gives $B_z \sim \exp(k^2 t_1) = \exp(k^2 S^{-1/3} t)$. Equation (35) then yields, near the X-point,

$$J_\perp(0, 0, t) = J_\perp(0, 0) \exp(k^2 S^{-1/3} t)\,, \tag{49}$$

which means that for $\tau_A \ll t \ll \tau_1$, the transverse current continues to grow exponentially with time, but with a characteristic (dimensional) growth rate $= \tau_R^{-1/3}\tau_A^{2/3}$. However, this time dependence, which is exponential in both the linear and ideal resistive phases, should be contrasted with the time dependence in Taylor's model (Hahm & Kulsrud 1985; Wang & Bhattacharjee 1992a) in which the current density at the separatrix grows algebraically with time.

5. THE NONLINEAR RESISTIVE PHASE

A necessary criterion for the validity of linear theory is that the magnetic island width be much smaller than the reconnection layer width. When the island width becomes comparable to the reconnection layer width, the reconnection process enters a nonlinear phase. In this phase, the dynamics conserves helicity, and the current sheet that develops in the linear phase is sustained. However, the spatial and temporal scales that characterize its nonlinear development are different from those in the linear phase (Waelbroeck 1989; Bhattacharjee & Wang 1991; Wang & Bhattacharjee 1992a, b). As discussed below, the reconnection dynamics in this phase is qualitatively similar to that in the Sweet-Parker model of magnetic reconnection (Sweet 1958; Parker 1957).

To investigate the nonlinear phase, it is useful to make the coordinate transformation $(x, y) \to (\psi, l_\perp)$ where $l_\perp$ is the length along a $B_\perp$ line measured from the point with coordinates $x = y$. (This applies to Fig. 2.) We assume that the inequalities

$$\frac{\partial}{\partial l_\perp} \ll \sqrt{\psi}\,\frac{\partial}{\partial \psi}\,, \quad \frac{\partial}{\partial t} \ll \frac{d\psi}{dt}\frac{\partial}{\partial \psi}\,, \tag{50}$$

hold. The solutions to the governing equations are then self-similar. Specifically, if $x \to \alpha x$ and $y \to \alpha y$, then $\boldsymbol{B}_\perp \to \alpha \boldsymbol{B}_\perp$ and $\psi \to \alpha^2 \psi$. (This scaling property is exactly satisfied by the exterior region solution for $\beta = 1$, given by Finn & Lau.) Near the X-point, $B_\perp = (2\psi)^{1/2}$.

We now describe a slight variant of the Sweet-Parker model. From equation (4), we get

$$d\psi/dt \simeq S^{-1} B_\perp/\delta\,, \tag{51}$$

where δ is the width of the current sheet. Outside the current sheet, at the inner limit of the exterior region, we have

$$\frac{d\psi}{dt} \simeq \boldsymbol{v} \cdot \nabla\psi = v_\psi B_\perp\,. \tag{52}$$

From the equation of continuity for an incompressible plasma, we get

$$v_\psi L = \delta v_{l_\perp}\,, \tag{53}$$

where L is length and δ is the width of the current sheet. From the ideal Ohm's law which holds in the exterior region, we have

$$v_{l_\perp} \sim B_\perp\,. \tag{54}$$

Using equations (53) and (54), we get

$$v_\psi \sim \delta B_\perp/L\,, \tag{55}$$

which implies, by equations (51) and (52), that

$$\delta \sim \frac{1}{S^{1/2}}\left(\frac{L}{B_\perp}\right)^{1/2}\,. \tag{56}$$

From equations (51) and (56), we then obtain

$$\frac{d\psi}{dt} \simeq \frac{1}{S^{1/2}}\frac{B_\perp^{3/2}}{L^{1/2}} \sim \frac{1}{S^{1/2}}\psi^{3/4}\,. \tag{60}$$

Integrating equation (60) with respect to time, we get

$$\psi \sim (S^{-1/2} t)^4\,. \tag{61}$$

Since $B_z = 0$ at $t = 0$, it follows from equation (5) that $B_z \simeq K_0 \psi$, where K_0 is a slowly varying function of $l_\perp$ and t. Therefore,

$$J_\perp \sim \frac{B_z}{\delta} \sim \frac{\psi^{5/4}}{\eta^{1/2}}\,, \tag{62}$$

which for $t \sim \tau_{SP} \equiv \tau_A^{1/2}\tau_R^{1/2}$, yields a current sheet of amplitude

$$J_\perp(0, 0, t \sim \tau_{SP}) \sim \frac{\tau_A^{5/2}}{\eta^{1/2}}\left(\frac{4\pi a^2}{c^2}\right)^{5/2} .$$

Thus, the current sheet has a width proportional to $S^{-1/2}$, and reconnection occurs on the characteristic Sweet-Parker timescale $S^{1/2}$ (which is $\tau_A^{1/2}\tau_R^{1/2}$ in actual time). We note that these spatial and temporal scales are very similar to those for the nonlinear $m = 1$ kink-tearing instability in a cylindrical loop (Strauss & Otani 1988; Waelbroeck 1989; Biskamp 1991). Following the discussion given elsewhere (Waelbroeck 1989; Bhattacharjee & Wang 1991), it can be shown from the constraint of helicity conservation that the X-point O is deformed to form Y-points.

It is interesting to note that Parker (1987) (see Appendix B of his paper) has discussed the problem of an inviscid incompressible fluid that is squeezed out from the region between two neighboring surfaces. Parker claims that if the two surfaces are mutually convex, they may be brought into contact in finite time whereas plane surfaces require infinite time. From this claim, it may be expected that if two surfaces that are initially convex change in time to planar surfaces, the rate of fluid expulsion should slow down. Although Parker's considerations do not directly apply to our model, there is a qualitative analogy: as the initial configuration near the separatrix flattens out in time from an X-point geometry to a Y-point geometry, the rate of reconnection as well as the growth of the current singularity slows down.

6. SUMMARY

A principal contribution of this paper is the *dynamical* description of current sheet formation and reconnection in two-dimensional coronal loops with X-type neutral lines. Previous considerations of this model have been essentially concerned with calculations of magnetostatic equilibria. Though those considerations are important because the system tends to evolve through a sequence of equilibria in the exterior region, it is equally important to consider what occurs in the interior region localized near the separatrix. The results presented in this paper demonstrate if we start from a smooth initial state, a true magnetostatic equilibrium is not established everywhere in the plasma in finite time. We show that the amplitude of the current sheet at the separatrix does tend to blow up while its width tends to shrink to zero, but this is merely a tendency and is not realized in practice. Thus, neither a true singularity in the current density nor a divergence in the footpoint displacement is dynamically accessible.

For the present model, we are able to obtain specific dynamical results within the framework of ideal as well as resistive MHD. In particular, we have given some exact solutions of the linearized ideal MHD equations which are valid on the Alfvén timescale. We find that the transverse current sheet amplitude grows as exp (t/τ_A) while its width shrinks as exp $(-t/\tau_A)$. The effects of a small but nonzero resistivity become important subsequently, and on the timescale $\tau_1 \equiv \tau_A^{2/3}\tau_R^{1/3}$ the current sheet grows as exp (t/τ_1). Following this linear resistive phase, the plasma enters the (nonlinear) Sweet-Parker regime, in which the transverse current density grown algebraically with time. Eventually, this current sheet smooths out on the diffusion timescale (τ_R).

We hope that the calculations given in this paper will put to rest some of the controversy that has dominated the discussion on current sheet formation in coronal plasmas. We address, in particular, the doubts raised by some computational studies on the validity of the current sheet picture. These computational studies demonstrate that though there are strong concentrations of current density, the current density does not lose its smoothness within the time of computation. For the analytical model considered in this paper, we show that this is indeed what should occur, and the inability to find δ-function current singularities in finite computing time is not due to a lack of numerical resolution. But even though the current sheet amplitude does not become infinitely large in finite time in the present model, it is sufficiently large that it causes intense coronal heating, as originally envisioned by Parker (1972). Moreover, the heating in the presence of X-type neutral lines appears to be an order of magnitude larger than in the straight cylindrical loop model considered in our earlier work (Wang & Bhattacharjee 1992b). If P_0 is the heating power for a loop without current sheets, and P the heating power in the presence of current sheets at X-type neutral lines, then, using equations (60), (61), and (63), of Wang & Bhattacharjee (1992b), we estimate that

$$\frac{P}{P_0} \sim \frac{\eta J_\perp^2 \delta a L}{\eta J_0^2 a^2 L} \sim S^{1/2} \sim 10^5 , \tag{64}$$

which is an order of magnitude higher than obtained before. Taking $P_0 \sim (10^{10}$–$10^{11})$ W, $a \sim 10^8$ cm, equation (64) give the energy flux $P/a^2 \sim (10^3$–$10^4)$ W m^{-2}. This heating can account for thermal energy balance in all quiet and many active loops.

This research is supported by the National Science Foundation under grant ATM 91-00513 and the Air Force Office of Scientific Research under grant F49620-93-1-0071.

APPENDIX

We consider the time evolution of a coronal plasma with power-law singularities in the initial state. Without loss of generality, we take

$$\psi_0 = (xy)^\beta , \qquad \beta < 1 . \tag{A1}$$

Repeating the procedure in § 3, we get

$$\frac{\partial B_z}{\partial t} = \beta\psi_0^z\left(y\frac{\partial v_z}{\partial y} - \frac{\partial v_z}{\partial x}\right), \tag{A2}$$

$$\frac{\partial v_z}{\partial t} = \beta\psi_0^z\left(y\frac{\partial B_z}{\partial y} - x\frac{\partial B_z}{\partial x}\right), \tag{A3}$$

where $\alpha = 1 - 1/\beta < 0$. If we define a new variable $\tau = \beta\psi_0^{\alpha} t$, equations (A2) and (A3) reduce to equations (20) and (21), respectively, except that the variable t is replaced by τ. It then follows that

$$J_{\perp} \simeq J_0 \exp(\beta\psi_0^{\alpha} t) . \tag{A4}$$

As $\psi_0 \to 0$, $\psi_0^{\alpha} \to \infty$ and hence, we obtain a current singularity in finite time at the separatrix where $\psi_0 = 0$. We note, however, that this finite-time singularity is an artifact of the initial condition (A2) which is itself singular. The dynamical evolution takes a singular initial state (with a power-law current singularity) to a more singular state (with an exponential current singularity) in finite time. If, on the other hand, we start from a smooth initial state (with $\beta = 1$), it is possible to approach (exponentially in time) but never truly access the singular equilibria with $\beta < 1$ in finite time.

REFERENCES

Aly, J. J. 1987, in Proc. Workshop on Interstellar Magnetic Fields, ed. R. Beck & R. Gräve (Berlin: Springer), 240
Antiochos, S. K. 1987, ApJ, 312, 886
Bhattacharjee, A., & Wang, X. 1991, ApJ, 372, 321
Biskamp, D. 1991, Phys. Fluids B, 3, 3353
Finn, J. M., & Y.-T. Lau. 1991, Phys. Fluids B, 3, 2675
Grad, H., Hu, P. N., & Stevens, D. C. 1975, Proc. Natl. Acad. Sci., 72, 3789
Greene, J. M. 1988, J. Geophys. Res., 93, 8583
Hahm, T. S., & Kulsrud, R. M. 1985, Phys. Fluids, 28, 2412
Jensen, T. H. 1989, ApJ, 343, 507
Karpen, J. T., Antiochos, S. K., & DeVore, C. R. 1990, ApJ, 356, L67
Longcope, D. W., & Sudan, R. N. 1992, ApJ, 384, 305
Low, B. C., & Wolfson, R. 1988, ApJ, 324, 579
Moffatt, H. K. 1987, in Advances in Turbulence, ed. G. Comte-Bellott & J. Mathieu (Berlin: Springer), 228
Parker, E. N. 1957, J. Geophys. Res., 62, 507
———. 1972, ApJ, 174, 499
Parker, E. N. 1987, ApJ, 318, 876
Rosenbluth, M. N., Dagazian, R. Y., & Rutherford, P. H. 1973, Phys. Fluids, 16, 1984
Strauss, H. R., & Otani, N. F. 1988, ApJ, 326, 418
Sweet, P. A. 1958, Electromagnetic Phenomena in Cosmical Physics (New York: Cambridge Univ. Press), 123
Syrovatsky, S. I. 1981, ARA&A, 19, 163
van Ballegooijen, A. A. 1985, ApJ, 298, 421
Vekstein, G. E., & Priest, E. R. 1992, ApJ, 384, 333
Waelbroeck, F. L. 1989, Phys. Fluids B, 1, 2372
Wang, X., & Bhattacharjee, A. 1992a, Phys. Fluids B, 4, 1795
———. 1992b, ApJ, 401, 371
Zweibel, E. G., & Li, H.-S. 1987, ApJ, 312, 423
Zweibel, E. G., & Proctor, M. R. E. 1990, in Topological Fluid Mechanics, ed. H. K. Moffat (Cambridge: Cambridge Univ. Press), 187
Zwingmann, W., Schindler, K., & Birn, J. 1985, Sol. Phys., 99, 133

Dynamics of current sheet formation and reconnection in two-dimensional coronal loops

Z. W. Ma, C. S. Ng, Xiaogang Wang, and A. Bhattacharjee
Department of Physics and Astronomy, University of Iowa, Iowa City, Iowa 52242

(Received 11 January 1995; accepted 27 April 1995)

Current sheet formation and magnetic reconnection in a two-dimensional coronal loop with an X-type neutral line are simulated numerically using compressible, resistive magnetohydrodynamic equations. Numerical results in the linear and nonlinear regimes are shown to be in good agreement with a recent analytical theory [X. Wang and A. Bhattacharjee, Astrophys. J. **420**, 415 (1994)]. The topological constraint imposed by helicity-conserving reconnection is discussed. It is found numerically that helicity-conserving reconnection causes the initial X-point structure of the loop to change to Y points, with current sheets at the separatrices encompassing the Y points. Implications for observations are discussed. © *1995 American Institute of Physics.*

I. INTRODUCTION

Two-dimensional (2-D) coronal loops with X-type neutral lines are of considerable interest from the standpoint of theory as well as observations. Such configurations are analytically and computationally tractable, and can represent with acceptable realism many observed features of coronal loops, reconstructed from X-ray images and magnetograms. In particular, it is widely believed that a study of such configurations can help elucidate the fundamental physical processes underlying a multitude of solar phenomena such as nanoflares, microflares, or x-ray bright points.

Several analytical and computational studies[1–9] of two-dimensional coronal loops with X-type neutral lines have been carried out, and have dealt with the issue of the existence of magnetostatic equilibria with current sheets. The general question that is addressed in these studies is the following: if we begin with an initial state that is a smooth solution of the magnetostatic equilibrium equations, is it possible, by means of smooth photospheric footpoint motions, to obtain neighboring equilibria with current sheets, i.e., singular current densities that result in jump discontinuities in the magnetic field? The existence of current sheets in the context of solar (and other astrophysical) plasmas has a crucial bearing on the coronal heating problem, as first pointed out by Parker,[10] who has given a comprehensive discussion of this problem in a recent monograph.[11]

The two-dimensional coronal loop with an X-type neutral line is a configuration in which there is no doubt regarding the tendency of the system to develop current sheets. This is because this system has well-defined separatrices, where the current density will tend to be singular for a large class of footpoint displacements. However, there has been some debate on the precise nature of the current singularities and the footpoint displacements that produce them. Finn and Lau[8] have argued that some equilibrium current-sheet solutions discussed in the literature are characterized by footpoint displacements that have logarithmic singularities[12,13] near the separatrix, and are therefore not admissible as a class of smooth, physical, footpoint displacements. To avoid these logarithmic singularities, Finn and Lau (as well as Vekstein and Priest[9]) have constructed scale-invariant equilibrium solutions that have smooth footpoint displacements and fractional power-law current singularities near the separatrix.

In all the analytical studies mentioned above, the focus is on the nature of the magnetostatic equilibrium solutions. The assumption is made implicitly that the coronal plasma will adjust to photospheric footpoint motion in order to maintain magnetostatic equilibrium, but no attempt is made to determine how rapidly in time the adjustment actually occurs. It should be emphasized that this is not merely an academic question. Observations of microflares and x-ray bright points in the solar corona exhibit a whole range of temporal dependencies from burstiness to relative quiescence, and the burden is on theory not only to provide a sequence of neighboring equilibria, but also to account for the time dependence of the growth and decay of an event.

Wang and Bhattacharjee[14] (hereafter referred to as WB) have investigated analytically the time evolution of the magnetic and velocity fields in a two-dimensional coronal loop with an X-type neutral line using the incompressible resistive magnetohydrodynamic (MHD) equations. The initial state in WB is a vacuum state with a hyperbolic X point, shown schematically in Fig. 1. The coronal loop lies in the x-y plane, with the footpoints of the magnetic field lines intersecting the photosphere, which is taken to be a plane perpendicular to the plane of the paper. The separatrices O_1O_1' and O_2O_2' cross at O, which is the point of intersection of the X-type neutral line with the x-y plane. As this state is driven by smooth photospheric footpoint motion, it tends to develop current sheets at the separatrices. Using the ideal MHD equations, WB derive an exact solution for the linear regime and show that the transverse current sheet $J_\perp$ (i.e., the current density perpendicular to $\hat{z}$) grows exponentially with time. Within a few characteristic Alfvén times, the effect of a small but finite resistivity slows down the rate of development of the current sheet. Subsequently, assuming that the footpoint motions at the boundary drive the system to a nonlinear regime, WB show that $J_\perp$ saturates, and that the longitudinal current density J_z grows algebraically with time.

One of the principal goals of this paper is to test the analytical results of WB (and the generalizations thereof, given in this paper) by numerical simulation of the full re-

 1070-664X/95/2(8)/3184/10/$6.00 © 1995 American Institute of Physics

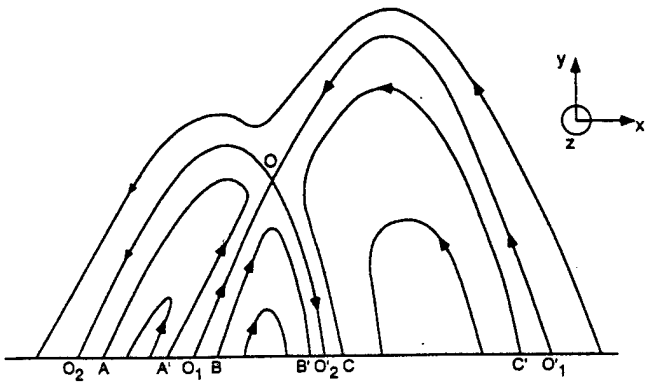

FIG. 1. Schematic picture of a two-dimensional coronal loop with an X-type neutral line that intersects the x-y plane of the loop at the point O.

sistive MHD equations, including the effect of plasma compressibility. The numerical results on the time development of the current sheet are found to be in good agreement with the analytical predictions, not only in the linear regime where the analytical solutions are exact, but also in the nonlinear regime in which the analytical results are derived from some strong assumptions. We also demonstrate that when the reconnection dynamics forced by the footpoint motion is helicity conserving,[15,16] the X-point structure in the initial state transforms to Y points.[17]

The following is a plan of this paper. In Sec. II, we generalize the analytical solutions of WB based on the incompressible MHD equations, including the effect of viscosity. In Sec. III, we discuss numerical results from the fully compressible code, and compare these results with the analytical predictions of WB. In Sec. IV, we present numerical results on the helicity conservation constraint. We conclude in Sec. V with a summary and a discussion of the implications of our results for observations.

II. SOME EXACT SOLUTIONS OF INCOMPRESSIBLE MHD

For configurations with translation symmetry along z, the magnetic field $\mathbf{B}$ can be represented as

$$\mathbf{B}=\hat{\mathbf{z}}\times\nabla\psi+B_z\hat{\mathbf{z}}, \tag{1}$$

where ψ is a flux function. An incompressible flow field $\mathbf{v}$ with the same symmetry can also be represented as

$$\mathbf{v}=\hat{\mathbf{z}}\times\nabla\phi+v_z\hat{\mathbf{z}}, \tag{2}$$

where ϕ is a streamfunction. The 2-D incompressible MHD equations for the four dependent variables, B_z, v_z, ψ, and ϕ can then be written (in cgs units) as

$$\frac{\partial B_z}{\partial t}+(\phi,B_z)=(\psi,v_z)+\frac{\eta c^2}{4\pi}\nabla^2 B_z, \tag{3}$$

$$\frac{\partial v_z}{\partial t}+(\phi,v_z)=\frac{1}{4\pi\rho}(\psi,B_z)+\frac{\nu}{\rho}\nabla^2 v_z, \tag{4}$$

$$\frac{\partial\psi}{\partial t}+(\phi,\psi)=\frac{\eta c^2}{4\pi}\nabla^2\psi, \tag{5}$$

$$\frac{\partial\omega}{\partial t}+(\phi,\omega)=\frac{1}{c\rho}(\psi,J_z)+\frac{\nu}{\rho}\nabla^2\omega, \tag{6}$$

where $J_z=(c/4\pi)\nabla^2\psi$ is the longitudinal (or z component of the) current density, $\omega=\nabla^2\phi$ is the z component of the vorticity, ρ is the (constant) density of the plasma, and $(f,g)=\hat{\mathbf{z}}\cdot\nabla f\times\nabla g$. The initial static equilibrium is taken to be a vacuum magnetic field described by

$$\psi_0=b_0xy/a, \tag{7}$$

where b_0 and a are constants, and the origin of the coordinate system is taken to coincide with point O in Fig. 1. In this initial state, $J_{z0}=0$ and remains so for all times for continuous footpoint displacements in the z direction. As shown by WB and, more generally here, infinite-time current singularities develop in the transverse current density $\mathbf{J}_\perp(=\hat{\mathbf{x}}\,\partial B_z/\partial y-\hat{\mathbf{y}}\,\partial B_z/\partial x)$ in the ideal limit.

Equations (1)–(6) can be cast in dimensionless form by redefining $\mathbf{B}/b_0\to\mathbf{B}$, $\mathbf{x}/a\to\mathbf{x}$, $t/\tau_A\to t$, $\psi/b_0a\to\psi$, $\phi/v_Aa\to\phi$, and $\omega\tau_A\to\omega$, where $\tau_A=a/v_A$ and $v_A=b_0/(4\pi\rho)^{1/2}$ is the Alfvén speed. The dimensionless equations are

$$\frac{\partial B_z}{\partial t}+(\phi,B_z)=(\psi,v_z)+\frac{1}{S}\nabla^2 B_z, \tag{8}$$

$$\frac{\partial v_z}{\partial t}+(\phi,v_z)=(\psi,B_z)+\frac{1}{R}\nabla^2 v_z, \tag{9}$$

$$\frac{\partial\psi}{\partial t}+(\phi,\psi)=\frac{1}{S}\nabla^2\psi, \tag{10}$$

$$\frac{\partial\omega}{\partial t}+(\phi,\omega)=(\psi,J_z)+\frac{1}{R}\nabla^2\omega, \tag{11}$$

where $S\equiv\tau_R/\tau_A$ is the Lundquist number, with $\tau_R\equiv 4\pi a^2/\eta c^2$, and $R\equiv\tau_\nu/\tau_A$ is the Reynolds number, with $\tau_\nu\equiv\rho a^2/\nu$.

It is clear by inspection of Eqs. (5) and (6) that the time evolution of ψ and ϕ is independent of B_z and v_z. This is a direct consequence of the assumption of incompressibility in two dimensions. Therefore, it is possible to solve for B_z and v_z with fixed ψ and ϕ, using suitable boundary conditions. In the ideal case, WB obtained a simple class of exact solutions of these equations. Here we give a general method for obtaining exact, time-dependent solutions of Eqs. (8) and (9), including the effects of dissipation in the special case $S=R$, with $\psi=\psi_0$ and $\phi=\phi_0=0$. (Note that the initial state in our study is a vacuum state that is qualitatively different from the stressed initial state considered in Refs. 18–20.)

Following WB, we define the Elsässer variables $f=B_z+v_z$, $g=B_z-v_z$. Equations (8) and (9) can then be combined to give

$$\frac{\partial f}{\partial t}=y\frac{\partial f}{\partial y}-x\frac{\partial f}{\partial x}+\frac{1}{S}\left(\frac{\partial^2 f}{\partial x^2}+\frac{\partial^2 f}{\partial y^2}\right) \tag{12}$$

and

$$\frac{\partial g}{\partial t}=-y\frac{\partial g}{\partial y}+x\frac{\partial g}{\partial x}+\frac{1}{S}\left(\frac{\partial^2 g}{\partial x^2}+\frac{\partial^2 g}{\partial y^2}\right), \tag{13}$$

which can be solved, subject to the initial conditions

$$f_0(x,y)\equiv f(x,y,t=0), \tag{14}$$

$$g_0(x,y)\equiv g(x,y,t=0). \tag{15}$$

Note that Eqs. (12) and (13) are linear. We now introduce the transformation

$$\xi=\lambda(t)y\equiv e^t y, \tag{16}$$

$$\zeta=x/\lambda(t), \tag{17}$$

$$T=\int_0^t \lambda^2(t')dt'=\left(\frac{1}{2}\right)(e^{2t}-1), \tag{18}$$

whence Eq. (12) becomes

$$\frac{\partial F}{\partial T}=\frac{1}{S}\left(\frac{\partial^2 F}{\partial \xi^2}+\frac{1}{\lambda^4}\frac{\partial^2 F}{\partial \zeta^2}\right), \tag{19}$$

where $F(\xi,\eta,T)=f(x,y,t)$. Setting $S=\infty$, we have the ideal solution $\partial F/\partial T=0$, which implies that

$$F(\xi,\zeta,T)=F(\xi,\zeta,T=0)=f_0(\xi,\zeta)=f_0(ye^t,xe^{-t}). \tag{20}$$

Equation (19) can be solved exactly using a Green's function. The solution is

$$F(\xi,\zeta,T)=\frac{S\lambda^2}{4\pi T}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}d\xi'\,d\zeta'\,f_0(\xi',\zeta') \times\exp\left(-\frac{S}{4T}[(\xi-\xi')^2+\lambda^4(\zeta-\zeta')^2]\right), \tag{21}$$

which gives

$$f(x,y,t)=\frac{S}{2\pi(1-e^{-2t})}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}dx'\,dy'\,f_0(x',y') \times\exp\left(-\frac{S}{2(e^{2t}-1)}[(ye^t-y')^2 +e^{4t}(xe^{-t}-x')^2]\right). \tag{22}$$

Equation (22) satisfies the initial condition (14). Equation (13), subject to the initial condition (15), can be solved in a similar manner. The solution is

$$g(x,y,t)=\frac{S}{2\pi(1-e^{-2t})}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}dx'\,dy'\,g_0(x',y') \times\exp\left(-\frac{S}{2(e^{2t}-1)}[(xe^t-x')^2 +e^{4t}(ye^{-t}-y')^2]\right). \tag{23}$$

From the functions f and g, we obtain

$$B_z(x,y,t)=\tfrac{1}{2}(f+g) \tag{24}$$

and

$$v_z(x,y,t)=\tfrac{1}{2}(f-g), \tag{25}$$

with the initial conditions

$$B_z(x,y,0)=\tfrac{1}{2}(f_0+g_0) \tag{26}$$

and

$$v_z(x,y,0)=\tfrac{1}{2}(f_0-g_0). \tag{27}$$

The exact solution obtained by WB from the linearized ideal equations is a special case of Eqs. (22) and (23). From the linearized resistive equations without viscosity, WB also obtained an asymptotic solution for a single Fourier component that holds in the restricted time domain $1\ll t\ll S^{-1/3}$. In order to compare this asymptotic solution (in its domain of validity) with the exact solution obtained here, it will be necessary to complete the asymptotic solution by including the effect of viscosity and calculate the inverse Fourier–Laplace transform, after summing over all Fourier components. Though the methodology is well understood,[21,22] there seems little point in pursuing it, now that we have the exact solutions.

Given initial conditions (14) and (15), we can easily carry out the integrals (22) and (23) (numerically, if necessary). In the Appendix, we discuss a special example for which one can solve Eqs. (12) and (13) analytically in closed form. Such exact, self-consistent solutions presented above allow us to settle some points of principle regarding the spatiotemporal structure of current sheets. These solutions are of interest because they exhibit current sheets with ψ frozen in the initial state (7) for all times, despite the presence of resistivity. [The resistive diffusion term on the right-hand side of Eq. (5) vanishes because $J_z=0$, although $\eta\neq 0$.] However, these solutions cannot be regarded as generic because they are essentially linear for all times and the convective nonlinearities play no role in their evolution. A change in boundary conditions (such as the imposition of footpoint motions in the $x-y$ plane that direct flows toward the separatrix and tend to pile up flux there) or the inclusion of plasma compressibility, will make these special solutions irrelevant because nonlinear reconnection will change qualitatively the dynamics of the system. Indeed, it is not difficult to see that the exact solutions eventually evolve to a state in which the assumption of incompressibility, on which they are predicated, breaks down. This is because the time-dependent, incompressible solutions with steady ψ and ϕ must satisfy the condition

$$\nabla(p+B_z^2/2)=0, \tag{28}$$

where p is the plasma pressure. As $t\to\infty$, $|\nabla B_z|$ becomes extremely large (infinity, in the ideal limit) at the separatrix, and it becomes impossible for the system to adjust to such large pressure gradients. Hence, even small departures from the incompressibility condition will lead to qualitative changes in the magnetic field geometry and flow pattern, and violate the condition of frozen-in ψ that holds for the exact, incompressible solutions. Magnetic reconnection will then intervene, changing qualitatively the nonlinear dynamics of the system.

III. COMPRESSIBLE DYNAMICS

The discussion at the end of Sec. II motivates us to include the effect of compressibility on the dynamics of the coronal loop. We have developed a computer code for this purpose, based on the following system of equations:

$$\frac{\partial \rho}{\partial t} = -\nabla\cdot(\rho \mathbf{v}), \tag{29}$$

$$\frac{\partial(\rho \mathbf{v})}{\partial t} = -\nabla\cdot\left[\rho \mathbf{v}\mathbf{v} + \left(p + \frac{B^2}{2}\right)\mathbf{I} - \mathbf{B}\mathbf{B}\right], \tag{30}$$

$$\frac{\partial \psi}{\partial t} = -\mathbf{v}\cdot\nabla\psi + \frac{1}{S}\nabla^2\psi, \tag{31}$$

$$\frac{\partial B_z}{\partial t} = -\nabla\cdot(B_z\mathbf{v}) + \mathbf{B}\cdot\nabla v_z + \frac{1}{S}\nabla^2 B_z, \tag{32}$$

$$\frac{\partial p}{\partial t} = -\nabla\cdot(p\mathbf{v}) - (\gamma-1)p\nabla\cdot\mathbf{v}. \tag{33}$$

Here $\mathbf{I}$ is the unit dyadic, $\gamma(=\frac{5}{3})$ is the ratio of the specific heats of the plasma, the variables $\mathbf{x}$, $\mathbf{v}$, t, $\mathbf{B}$, and ψ are normalized to be dimensionless in the same manner as in Sec. II, and the pressure p is made dimensionless by scaling it with $b_0^2/4\pi$.

The initial force-free equilibrium containing an X-type neutral line is obtained by solving Laplace's equation $\nabla^2\psi_0=0$ in the whole physical domain as a boundary-value problem. The flux function at the bottom boundary is chosen to be

$$\psi_0(x,y=0) = \begin{cases} \Psi_0 \sin[(2\pi a/L)(x_0-x)], & x_1<x<x_0, \\ \Psi_0 \sin[(2\pi a/L)(x_0+x)], & -x_0<x<-x_1, \\ 0, \quad \text{otherwise}, \end{cases} \tag{34}$$

where x_0 and x_1 are control parameters, chosen in this simulation to be equal to $0.7L/a$ and $0.2L/a$, respectively. The normalization constant Ψ_0 is chosen to make the dimensionless magnetic field $\mathbf{B}$ unity on the boundary. The flux function is set to zero at the other three boundaries. The initial field configuration, which is determined by a numerical solution of Laplace's equation, is shown in Fig. 2. With the initial velocity and B_z set equal to zero, we have a static equilibrium.

Equations (29)–(33) are solved simultaneously using a Runge–Kutta finite-differencing scheme that has an accuracy of fourth order in time and second order in space. Exploiting the symmetry of the initial conditions, the numerical simulation is carried out in half of the physical domain in the x-y plane, with $x\in(0,1)$ and $y\in(0,2)$. To reduce numerical error and save computer time, we employ a nonuniform mesh. Nonuniform meshing enables us to increase resolution near the separatrix (where the current sheet develops) so that, with a 161×241 array, we can resolve 0.002 in x and 0.004 in y.

At the lateral mirror boundary $x=0$, dependent physical variables $G(x,y)$ are grouped as symmetric, i.e., $G(x,y)=G(-x,y)$, or antisymmetric, i.e., $G(x,y)=-G(-x,y)$. The fields ρ, p, v_y, ψ, and B_z are symmetric, whereas v_x and v_y are antisymmetric. The other lateral boundary $x=1$ as well as the upper boundary $y=2$ are treated as open, i.e., all variables except ψ obey the condition $\partial G/\partial t=0$, and ψ obeys the condition $\partial\psi/\partial t=-\mathbf{v}\cdot\nabla\psi$. In order to reduce the amplitude of waves reflected from the boundary, a damping term is applied at the outermost grid points. A three-point differencing method is used to calculate the first derivative at the boundary, ensuring second-order spatial accuracy.

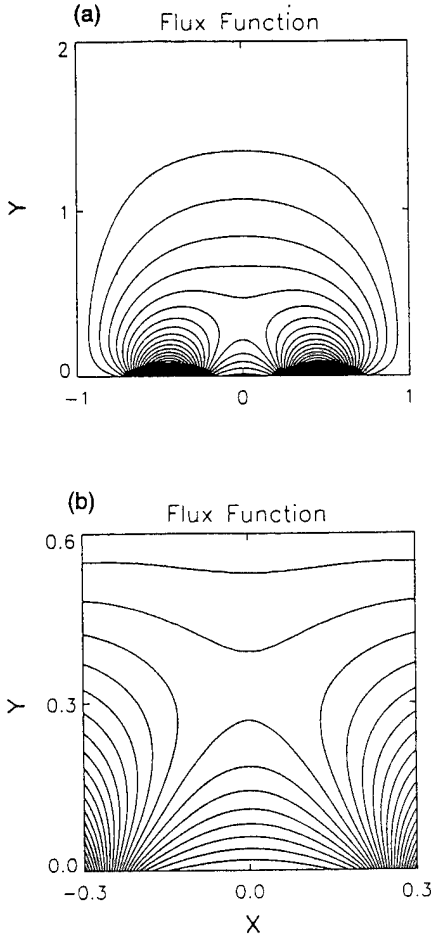

FIG. 2. Flux surfaces for a numerically generated vacuum-field solution of the 2-D loop. (b) is a magnified picture of the region around the X point in (a). This equilibrium solution is the initial condition for the compressible simulation.

The bottom boundary, assumed to simulate the photosphere, is treated as a material reservoir. All dependent variables on this boundary, except B_z and v_z, are held fixed at their initial values. Shearing motion of the footpoints along the z axis is prescribed according to the relation

$$v_z(x,y=0,t) = v_0 \tanh\left(\frac{t}{t_0}\right)\exp\left[-\left(\frac{x-x_0}{\Delta x}\right)^2\right], \tag{35}$$

where $x=x_0$ is the point at which the separatrix intersects the $y=0$ line. For the case studies discussed below, we have chosen $v_0=0.05$, $t_0=5$, and $\Delta x=0.05$. Since distances are normalized by a and time by τ_A, Eq. (35) describes smooth, coherent footpoint motions that build up to their maximum amplitude on a time scale that is nearly an order of magnitude larger than τ_A. The field B_z is obtained by direct numerical integration of Eq. (32), with one-sided differencing for derivatives with respect to y.

WB developed an analytical theory for the nonlinear regime, but they did not discuss the precise mechanisms by which the nonlinear regime may be realized. Footpoint motions in the x-y plane provide one way to realize the nonlin-

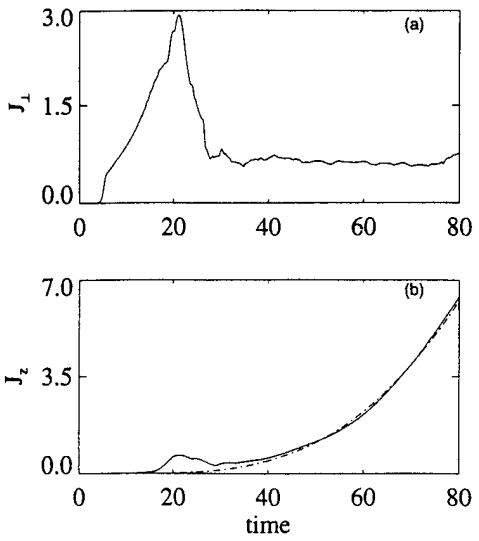

FIG. 3. Maximum amplitudes (a) $(\mathbf{J}_\perp)_{max}$ and (b) $(J_z)_{max}$ as a function of time from the compressible simulation with $S=2\times10^5$. The dashed line in (b) represents the analytically predicted t^3 behavior.

ear regime. Footpoint displacements in the z direction can also lead to the nonlinear regime, if plasma compressibility is invoked to couple ψ and ω with B_z and v_z. We pursue the second possibility here using our compressible code, and find that the results of the simulation on the time dependence of the current sheet development is in good accord with the analytical predictions of WB.

Figures 3(a) and 3(b) show the maximum amplitudes of $|\mathbf{J}_\perp|_{max}$ and $|J_z|_{max}$, respectively, for a typical run with $S=2\times10^5$. In the linear and the early nonlinear phase (up to $t\approx25$), the growth occurs primarily in $\mathbf{J}_\perp$ near the separatrix. In the linear phase, this growth is exponential in time, as shown by WB, and confirmed by the numerical results. Subsequently, due to the dynamical coupling caused by finite plasma compressibility, the system enters a nonlinear phase in which $\mathbf{J}_\perp$ decays and saturates, while J_z grows algebraically in time to large values near the separatrix. WB predicted that this growth can be described by the relation $\psi\sim(S^{-1/2}t)^4$, which implies that

$$J_z\sim\eta t^3-\eta^{-1/2}(t/\tau_{SP})^3, \tag{36}$$

where $\tau_{SP}\equiv(\tau_A\tau_R)^{1/2}$ is the characteristic Sweet–Parker time scale.[23] The dashed line in Fig. 3(b) is the analytically predicted t^3 behavior.

Figures 4(a) and 4(b) show the contour plots of J_z and ψ at $t=80$, respectively, for the run described in Fig. 3. Note the strong concentration of current density near the separatrix. The initial X-point structure of the separatrix has been altered to form Y points. This type of behavior has also been seen in the incompressible simulations of Biskamp[23] in a different geometry (with different boundary conditions). As explained in Sec. II, in the present geometry (Fig. 1) and with the boundary conditions (35), compressibility is crucial because it provides a mechanism for changing ψ by diffusion. Compressibility is known to play a similar role in numerical studies of magnetic arcade evolution (see, for instance, Biskamp and Welter[24] and other references therein).

IV. HELICITY-CONSERVING RECONNECTION AND Y POINTS

It has been suggested by Syrovatsky[17] that current sheets will be formed along Y points in a 2-D, quasi-ideal, quasistatic plasma. (Interested readers will find a picture of Y points in the earlier work of Sweet,[25] who seems to have been aware of, but did not prove, the possibility of Y points

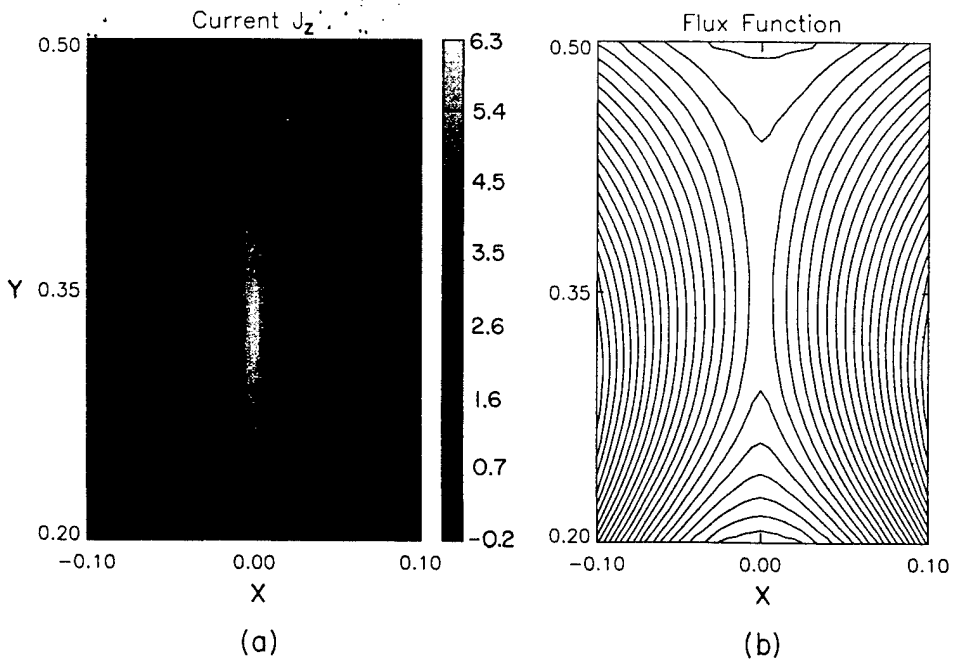

FIG. 4. Contour plots of (a) J_z and (b) ψ at $t=80$ for the compressible simulation with $S=2\times10^5$. Note the strong current sheet at the separatrix. The initial X-point structure, seen in Fig. 2, has changed to Y points.

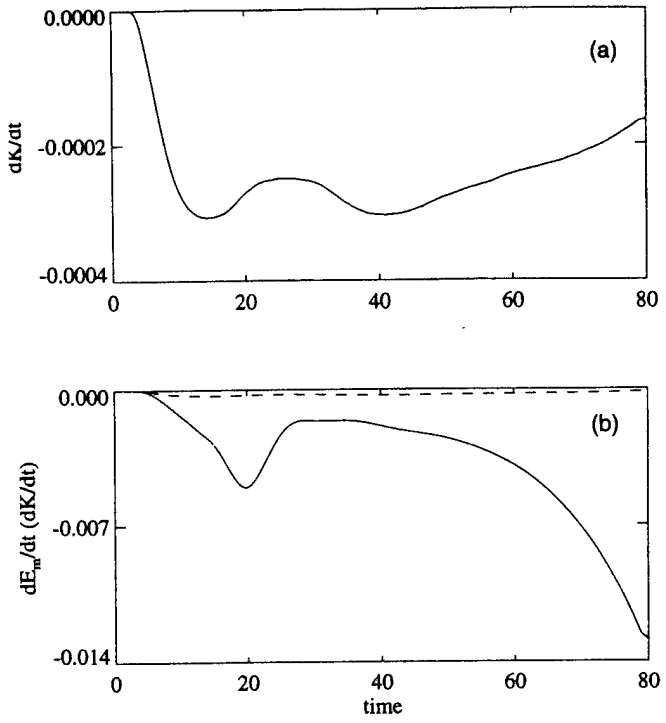

FIG. 5. Time rate of resistive decay of (a) magnetic helicity K and (b) magnetic energy E_m for the compressible simulation with $S=2\times10^5$. The rate of dissipation of energy is larger than that of helicity by nearly two orders of magnitude. The dashed line in (b) represents dK/dt.

in his steady-state reconnection model.) Further generalizations and refinements of Syrovatskii's 2-D results, derived using complex variable theory, have been given in Refs. 23 and 26–29. Although Syrovatskii's model is not fully self-consistent, its essential features have been confirmed by several self-consistent simulations in the regime of high Lundquist numbers.[23,29–31]

Here we present numerical evidence that helicity-conserving reconnection leads to the formation of Y points in 2-D coronal loops. It should be emphasized that helicity is not an exact invariant in the presence of diffusion, and so helicity-conserving reconnection refers to dynamics that conserves helicity much better than energy. It is also important to note that here we are considering a stronger form of the helicity conservation constraint than used by Taylor.[32] Taylor has discussed the self-organization of a strongly turbulent plasma in which it is not physically meaningful to speak of the topological identity of individual flux tubes. In a strongly turbulent system, Taylor has predicted that the total magnetic helicity is approximately conserved while the total magnetic energy decays rapidly. In contrast, what we have in mind is a much stronger form of the helicity conservation constraint, considered by Kadomtsev[33] and Bhattacharjee *et al.*,[34] who treat an unstable (but nonturbulent) plasma with well-defined magnetic surfaces nearly everywhere, in which magnetic helicity is conserved during "nonconstant-ψ" reconnection. Since there are, in principle, an infinite number of flux tubes in a plasma with good surfaces, there are an infinite number of constraints, whereas the Taylor relaxation theory keeps only one global constraint.[34] Using the form of the helicity conservation constraint developed in Refs. 33 and 34, and building on the mathematical framework developed by Rosenbluth *et al.*,[35] Waelbroeck[15] has shown that Y points develop in the nonlinear stage of a kink-tearing instability. Bhattacharjee and Wang[16] have adapted Waelbroeck's method to a "long-thin" coronal tube and shown that current sheets develop along the separatrix spanning Y points. [The "long–thin" approximation is motivated by the fact that coronal loops with very large aspect ratios ($\sim10^2$) are not uncommon.[36] The advantage of the approximation is that it makes the problem essentially two dimensional. Furthermore, the equilibrium is topologically sufficiently simple that it is possible to carry through an analytical calculation involving the helicity conservation constraint. However, such a calculation cannot be applied to all coronal tubes, because the simple notion of a 2-D separatrix does not carry over to a three-dimensional tube of finite length.]

In the present context, the law of helicity conservation means that the integral,

$$K=\int \mathbf{A}\cdot\mathbf{B}\,dx\,dy, \tag{37}$$

is (approximately) conserved for reconnecting flux tubes. Here $dx\,dy$ is the infinitesimal element of area in a section of a flux tube, bounded by two neighboring surfaces. It is easy to show that this is equivalent to the conservation of $\int dl/B$ for reconnecting flux tubes. Though simple in principle, it is difficult in practice to calculate K analytically for the complicated loop configurations of the type considered in this paper. It is easier to calculate the rate of resistive decay of helicity (and energy), because the dominant effect of re-

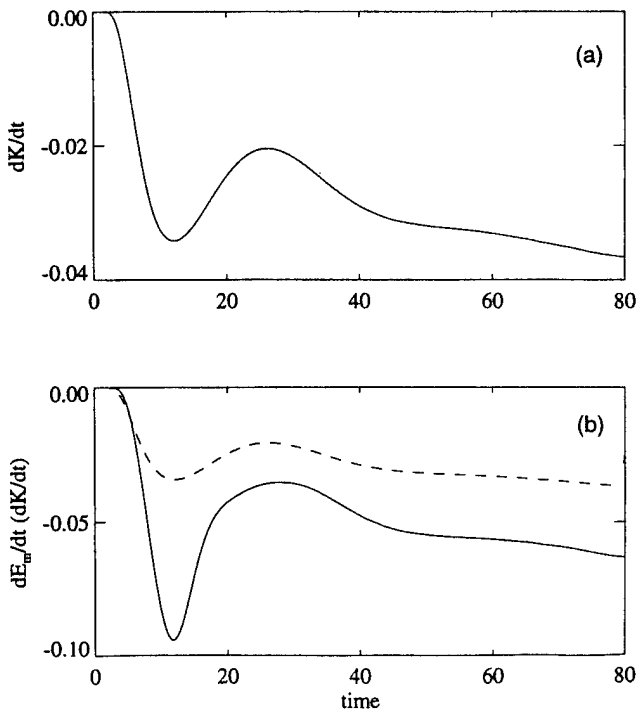

FIG. 6. Time rate of resistive decay of (a) magnetic helicity K and (b) magnetic energy E_m for the compressible simulation with $S=10^3$. The rate of dissipation of energy is comparable with that of helicity. The dashed line in (b) represents dK/dt.

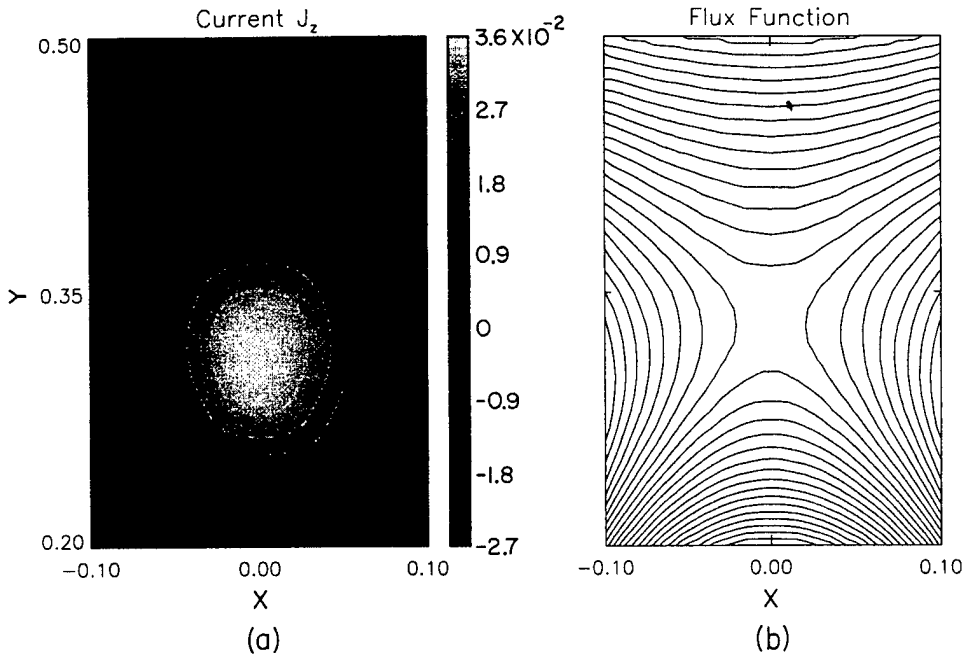

FIG. 7. Contour plots of (a) J_z and (b) ψ at $t=80$ for the compressible simulation with $S=10^3$. The amplitude of the current sheet is smaller, and it is less localized than it is in Fig. 4. The Y points are not well formed.

sistivity is localized near the separatrix where reconnection occurs. In Fig. 5, we show dK/dt and dE_m/dt (where $E_m \equiv \int dx\, dy\, B^2/2$) for the computer runs described in Figs. 3 and 4, with $S=2\times10^5$. The integrals are taken over a rectangular box that encompasses the reconnection layer. Though the magnetic helicity is not a constant [as indicated by the temporal variation in Fig. 5(a)], the change in helicity is approximately two orders of magnitude smaller than the change (loss) of energy. Hence, helicity can be regarded as approximately constant on the time scale magnetic energy decays. The numerical evidence supports the basic theoretical claim[14–16] that Y points are formed in helicity-conserving reconnection.

It is interesting to contrast the results of Fig. 5 with Fig. 6, which begins with the same initial condition, but uses a much lower value of $S(=10^3)$. In the latter case, the rates of resistive decay of helicity and energy are comparable, and helicity cannot be regarded as a good invariant while energy decays. Figure 7 shows the contour plots of ψ and J_z in this case, and it is obvious by inspection that the dynamics does not yield Y points. The current sheet J_z is also much weaker and spread out over a much broader diffusion layer. Figures 8(a) and 8(b) show the time evolution of $\mathbf{J}_\perp$ and J_z, respectively. The system evolves into the nonlinear phase very quickly, and the nonlinear growth of J_z is approximately linear in time (dashed curve), which is at variance with the analytical predictions and the numerical results given in Fig. 3. This underscores an issue that is central to discussions of current sheets and magnetic reconnection: low-S simulations ($\leqslant 10^3$) generally exhibit such large differences from high-S simulations ($\geqslant 10^5$) and analytical theories that it is difficult to extrapolate quantitative scalings obtained from low-S simulations to the high-S regimes that are believed to be observationally relevant for the solar corona.

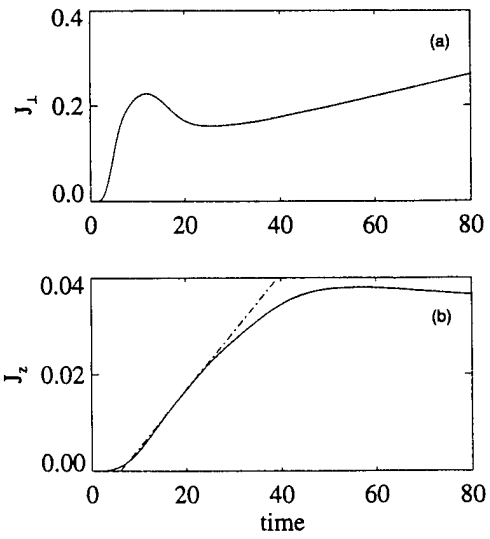

FIG. 8. Maximum amplitudes (a) $(\mathbf{J}_\perp)_{max}$ and (b) $(J_z)_{max}$ as a function of time from the compressible simulation with $S=10^3$. The dashed line in (b) is an approximate fit and shows $(\mathbf{J}_\perp)_{max}$ increasing linearly with time.

V. SUMMARY AND DISCUSSIONS

The results of this paper, when combined with those obtained earlier by Wang and Bhattacharjee,[14] provide a detailed description of the dynamics of current sheet formation in a two-dimensional coronal loop with an X-type neutral line. The emphasis in Ref. 14, as well as the present paper, is

on the precise time history of current sheets in such a loop beginning from a vacuum field, driven by smooth photospheric footpoint motion.

For compressible plasmas, we have presented numerical results on current sheet development in the linear and nonlinear stages of the dynamics. In the linear regime, on ideal time scales, current sheets tend to grow exponentially in $\mathbf{J}_\perp$. This linear, ideal regime is followed by a short-lived linear, resistive regime before the system passes into the nonlinear regime on the characteristic Sweet–Parker time scale $(\tau_A \tau_R)^{1/2}$. In the early nonlinear regime, $\mathbf{J}_\perp$ saturates while decays J_z begins to grow. The growth in J_z is algebraic in time, proportional to t^3, predicted by the analytical theory of WB and confirmed here by numerical solutions. It is found that the X-point structure of the separatrix in the initial state changes to Y points in helicity-conserving reconnection, as predicted by theory in the high-S regime.[14–16]

It is worthwhile to reflect on the physical significance of these results, and what they imply for observations. For smooth initial conditions, all the current singularities we find are infinite-time singularities in the ideal limit. At least in the context of the present geometry, this should put to rest the question as to whether a true current singularity can be seen in ideal simulations by extrapolating to zero mesh size. As discussed in Ref. 14, the answer is no, because current singularities do not occur in finite time, even in the absence of dissipation. The current sheet keeps growing until it attains an amplitude large enough that the effect of diffusion has to be included as an essential part of the dynamics. It is interesting to note that the current sheet J_z, which actually owes its growth to the presence of diffusion and magnetic reconnection, ends up throttling the nonlinear reconnection dynamics to yield algebraic growth in time.

From the point of view of observations, there are several issues that need to be considered. It is clear that current sheets of the type considered here can produce significant coronal heating. Simple estimates, assuming that such sheets have a filling factor of the order unity, have been given in Ref. 14 and indicate that the heating produced can account for quiet loops and some, but not all, active loops. Direct observational tests of the validity of current sheet models, as discussed in considerable depth by Parker in his recent monograph,[11] involve several features that are beyond the scope of present-day observational techniques. The small spatial scales of current sheets are well below the spatial resolution of the state of the art instruments. One has to rely, therefore, on the integrated effect of numerous transient events as it affects a measurable quantity, but the interpretation of the latter is often subject to the uncertainties inherent in the various transport processes in the corona.[37] For example, assuming that the corona is made up of many hundred loops that are heated by nanoflares, Cargill[38] has developed a model to determine the dependence of the emission measure on temperature and established good correspondence with observations at temperatures in the range $(10^5–10^6)$ K. But the modeling of the emission measure depends on parallel heat conduction, regarding which there are major uncertainties. In view of all this, as Cargill himself points out, one cannot regard agreement with emission measures in a restricted temperature range, as sufficient evidence to rule exclusively in favor of current sheet models, but it is a minimum requirement that any heating mechanism should satisfy. For a comparative discussion of the various heating mechanisms and how they measure up against observations, the reader is referred to the recent review by Cargill,[39] who has also given an extensive bibliography.

Some of the more effective diagnostics for nonlinear reconnection theories are the observations on the temporal variation of hard x-ray emission (>10 keV). Hard x rays are a reliable signature of prompt electrons, and if one assumes that MHD activity causes the acceleration of electrons, then one can look at the temporal variation of the x-ray emission as a qualitative predictor of the time dependence of the electric field generated by reconnection. Such data is abundant from solar flares,[40] much less so for microflares,[41] and nonexistent for nanoflares. However, if we proceed on the tentative premise that reconnection phenomena of different degrees of virulence are responsible for the various types of flares, then one can hope to identify certain patterns in the time dependence of the x-ray emission, even if they correspond to events that differ vastly in the amount of energy liberated. [It is the amount of energy liberated that distinguishes a typical flare ($\sim 10^{33}$ ergs) from a typical microflare ($\sim 10^{27}$ ergs) and a typical nanoflare ($\sim 10^{24}$ ergs).] For instance, it is well known that in a typical flare, the time dependence of the hard x-ray emission exhibits an impulsive phase in which the emission exhibits a sudden transition in time from the slower growth phase. There is a similar trend in the data on microflares[40] that exhibit x-ray spikes, each of which one may assume is due to an individual reconnection event.[11] If we accept this sudden onset as a recurrent pattern, then the burden is on reconnection theories to account, not only for the rapidity of the impulsive phase, but the time development of the entire process, which includes the sudden transition to the impulsive phase. This is where nonlinear reconnection models, such as the one discussed in the present paper, run into difficulties, because in such models the reconnection dynamics passes from a linear exponential phase to an algebraic phase in time, and this is too qualitatively gradual to account for the sudden change in the time derivative of the emission profile, as the impulsive phase is triggered. It is possible that this discrepancy may be cured by considering a different class of boundary conditions at the photosphere or by going beyond the resistive MHD model to include other collisionless effects in a generalized Ohm's law. In the context of the analogous problem of sawtooth trigger in tokamaks, Wang and Bhattacharjee[42] have recently shown that electron pressure gradient effects (in a generalized Ohm's law) can change the algebraic growth rates from nonlinear resistive MHD model to near-explosive growth. The exploration of this effect, as well as the dependencies on other types of footpoint motions on the boundary, is left to future work.

ACKNOWLEDGMENTS

This research is supported by the National Science Foundation Grant No. ATM 93-10157 and the Air Force Office of Scientific Research Grant No. F49620-93-1-0071.

APPENDIX: SOME EXACT INCOMPRESSIBLE SOLUTIONS

We consider some exact solutions of Eqs. (22) and (23). We take

$$B_z(x,y,t=0)=\nu_z(x,y,t=0)=B_0\ \mathrm{erf}[-k(y+y_0)], \tag{A1}$$

where erf denotes the error function, and B_0, k, and y_0 are constants, with k, $y_0>0$. In particular, we choose $ky_0\gg 1$, so that the initial perturbation of the separatrix is very small. The initial condition (A1) yields the exact, time-dependent solution,

$$B_z(x,y,t)=\nu_z(x,y,t)$$
$$=B_0\ \mathrm{erf}\left(-\frac{ke^t(y+y_0e^{-t})}{[1+(2k^2/S)(e^{2t}-1)]^{1/2}}\right). \tag{A2}$$

The transverse current density is given by

$$\mathbf{J}_\perp=-\frac{2B_0}{\sqrt{\pi}}\frac{ke^t}{[1+(2k^2/S)(e^{2t}-1)]^{1/2}}$$
$$\times\exp\left(-\frac{[ke^t(y+y_0e^{-t})]^2}{1+(2k^2/S)(e^{2t}-1)}\right)\hat{\mathbf{x}}. \tag{A3}$$

For $t<(\frac{1}{2})\ln(S/2k^2)$, we obtain the ideal solution,

$$B_z(x,y,t)=\nu_z(x,y,t)=B_0\ \mathrm{erf}[-ke^t(y+y_0e^{-t})], \tag{A4}$$

with the transverse current density

$$\mathbf{J}_\perp=-\frac{2kB_0e^t}{\sqrt{\pi}}\exp[-k^2e^{2t}(y+y_0e^{-t})^2]\hat{\mathbf{x}}. \tag{A5}$$

The right-hand side of Eq. (A5) tends to a δ function as $t\to\infty$. In other words, a current sheet is an infinite-time singularity of Eqs. (8)–(11) in the ideal limit.

For $t>(\frac{1}{2})\ln(S/2k^2)$, in the limit of large S, Eqs. (A3) and (A4) yield

$$B_z=B_0\ \mathrm{erf}\left(-\sqrt{\frac{S}{2}}(y+y_0e^{-t})\right) \tag{A6}$$

and

$$\mathbf{J}_\perp=-\sqrt{\frac{2S}{\pi}}B_0\exp\left(-\frac{S}{2}(y+y_0e^{-t})^2\right)\hat{\mathbf{x}}, \tag{A7}$$

respectively. As $t\to\infty$, Eqs. (A5) and (A6) describe a steady state in which there is a saturated current sheet at the separatrix, with an amplitude that increases as $S^{1/2}$ and a width that shrinks as $S^{-1/2}$.

For the general case $S\neq R$, although we have not found an exact, time-dependent solution, it can be shown that there exists a quasistationary solution to which the system may relax as $t\to\infty$. If the viscosity and resistivity are unequal, the equations for B_z and ν_z are given, respectively, by

$$\frac{\partial B_z}{\partial t}=y\frac{\partial\nu_z}{\partial y}-x\frac{\partial\nu_z}{\partial x}+\frac{1}{S}\left(\frac{\partial^2B_z}{\partial x^2}+\frac{\partial^2B_z}{\partial y^2}\right) \tag{A8}$$

and

$$\frac{\partial\nu_z}{\partial t}=y\frac{\partial B_z}{\partial y}-x\frac{\partial B_z}{\partial x}+\frac{1}{R}\left(\frac{\partial^2\nu_z}{\partial x^2}+\frac{\partial^2\nu_z}{\partial y^2}\right), \tag{A9}$$

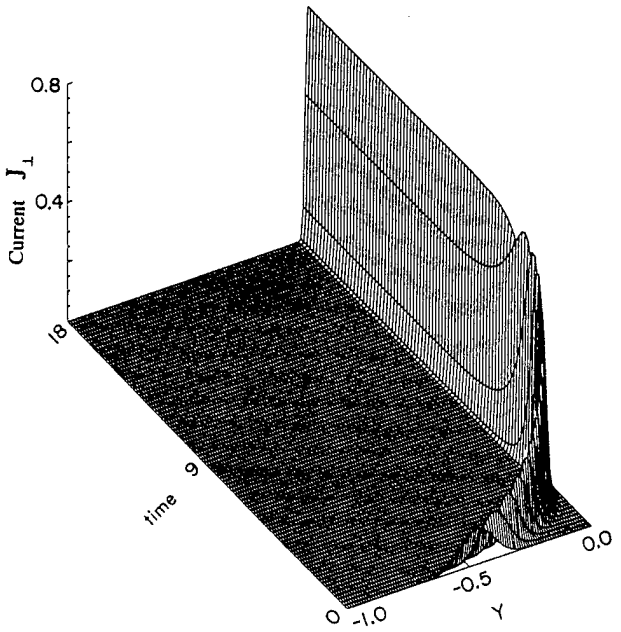

FIG. 9. The growth of a current sheet $\mathbf{J}_\perp$ in an incompressible, resistive plasma without reconnection from the smooth initial condition (A1), with $S=R=10^4$. The current profile saturates as $t\to\infty$.

respectively. We investigate quasistationary solutions of Eqs. (A8) and (A9) in the limit $t\to\infty$. Setting $\partial B_z/\partial t=\partial\nu_z/\partial t=0$, and defining $B'\equiv B_z/k$, where k is a constant, we obtain

$$y\frac{\partial\nu_z}{\partial y}-x\frac{\partial\nu_z}{\partial x}+\frac{k}{S}\left(\frac{\partial^2B'}{\partial x^2}+\frac{\partial^2B'}{\partial y^2}\right)=0 \tag{A10}$$

and

$$y\frac{\partial B'}{\partial y}-x\frac{\partial B'}{\partial x}+\frac{1}{kR}\left(\frac{\partial^2\nu_z}{\partial x^2}+\frac{\partial^2\nu_z}{\partial y^2}\right)=0. \tag{A11}$$

In order to make Eqs. (A10) and (A11) symmetric, we choose $k=\sqrt{S/R}$. Defining new Elsässer variables, $f'=B'+\nu_z$ and $g'=B'-\nu_z$, we obtain the quasistationary solutions,

$$g'=0 \tag{A12}$$

and

$$f'=B_0\left[1+\mathrm{erf}\left(-\sqrt{\frac{S}{2k}}y\right)\right]. \tag{A13}$$

Equations (A12) and (A13) yield

$$B_z=\frac{B_0}{2}\sqrt{\frac{S}{R}}\left\{1+\mathrm{erf}\left[-\left(\frac{SR}{4}\right)^{1/4}y\right]\right\}, \tag{A14}$$

which corresponds to a transverse current sheet with an amplitude directly proportional to $(S^3/R)^{1/4}$ and a width inversely proportional to $(SR)^{1/4}$.

We have developed a computer code that integrates Eqs. (8)–(11). This provides a check for the exact analytical solutions in the case $S=R$, and also determines the time development of the solutions in the cases $S\neq R$, which break the symmetry of Eqs. (12) and (13). In Fig. 9, we show the growth of the current sheet from the smooth initial state (A1)

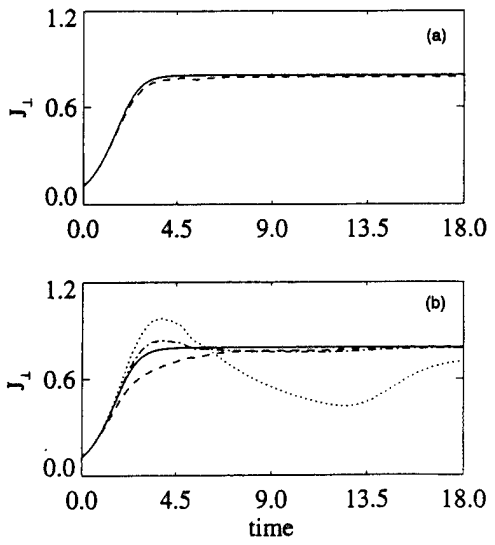

FIG. 10. (a) Comparison of the numerical (dashed line) and analytical (solid line) results for $(\mathbf{J}_\perp)_{max}$ as a function of time, with $S=R=10^4$. (b) Numerical results for $(\mathbf{J}_\perp)_{max}$ in three cases: $S=10^4$, $R=5\times10^3$(---); $S=10^4$, $R=2\times10^4$(-·-); and $S=10^4$, $R=\infty$(···), compared with the exact analytical solution (solid line) for $S=R=10^4$, presented in Figs. 9 and (a).

to saturation for a case with $S=R$. In Fig. 10(a), we show that the numerical (dashed line) and analytical (solid line) results for the maximum transverse current density amplitude ($|\mathbf{J}_\perp|_{max}$) are in good agreement for the parameters of Fig. 9. In Fig. 10(b), we present the numerical solutions for $|\mathbf{J}_\perp|_{max}$ in three cases with $S\neq R$, and compare them with the analytical result of Fig. 10(a). As may be expected, the time development in Fig. 10(b) does deviate from the exact analytical prediction for $S=R$, and especially so for the case of zero viscosity. In the presence of viscosity, the mechanical oscillations are gradually damped, and the system tends to a quasistationary state.

[1]J. J. Aly, in *Proceedings of the Workshop on Interstellar Magnetic Fields*, edited by R. Beck and R. Gräve (Springer-Verlag, Berlin, 1987), p. 240.
[2]H. K. Moffatt, in *Advances in Turbulence*, edited by G. Comte-Bellott and J. Mathieu (Springer-Verlag, Berlin, 1988), p. 228.
[3]B. C. Low and R. Wolfson, Astrophys. J. **324**, 579 (1988).
[4]T. H. Jensen, Astrophys. J. **343**, 507 (1989).
[5]J. T. Karpen, S. K. Antiochos, and C. R. Devore, Astrophys. J. **356**, L67 (1990).
[6]E. G. Zweibel and M. R. E. Proctor, in *Topological Fluid Mechanics*, edited by H. K. Moffatt (Cambridge University Press, Cambridge, 1990), p. 187.
[7]G. Vekstein, E. Priest, and T. Amari, Astron. Astrophys. **243**, 492 (1991).
[8]J. M. Finn and Y. T. Lau, Phys. Fluids B **3**, 2675 (1991).
[9]G. E. Vekstein and E. R. Priest, Astrophys. J. **384**, 333 (1992).
[10]E. N. Parker, Astrophys. J. **174**, 499 (1972).
[11]E. N. Parker, *Spontaneous Current Sheets in Magnetic Fields* (Oxford University Press, New York, 1994).
[12]H. Grad, P. N. Hu, and D. C. Stevens, Proc. Natl. Acad. Sci. USA **72**, 3789 (1975).
[13]W. Zwingmann, K. Schindler, and J. Birn, Sol. Phys. **99**, 133 (1985).
[14]X. Wang and A. Bhattacharjee, Astrophys. J. **420**, 415 (1994).
[15]F. L. Waelbroeck, Phys. Fluids B **1**, 2372 (1989).
[16]A. Bhattacharjee and X. Wang, Astrophys. J. **372**, 321 (1991).
[17]S. I. Syrovatsky, Sov. Phys. JETP **33**, 933 (1971).
[18]I. J. Craig and A. N. McClymont, Astrophys. J. **393**, 385 (1992).
[19]A. B. Hassam, Astrophys. J. **399**, 159 (1992).
[20]L. Ofman, P. J. Morrison, and R. S. Steinolfson, Astrophys. J. **417**, 748 (1993).
[21]T. S. Hahm and R. M. Kulsrud, Phys. Fluids **28**, 2412 (1985).
[22]X. Wang and A. Bhattacharjee, Phys. Fluids B **4**, 1795 (1992).
[23]D. Biskamp, Phys. Fluids **29**, 1520 (1986).
[24]D. Biskamp and H. Welter, Sol. Phys. **120**, 49 (1989).
[25]P. A. Sweet, *Electromagnetic Phenomena in Plasma Physics* (Cambridge University Press, Cambridge, 1958), p. 123; E. N. Parker, J. Geophys. Res. **62**, 507 (1957).
[26]E. R. Priest and M. R. Raadu, Sol. Phys. **43**, 177 (1975).
[27]Y. Q. Hu and B. C. Low, Sol. Phys. **81**, 107 (1982).
[28]T. Amari and J. J. Aly, Astron. Astrophys. **227**, 628 (1990).
[29]D. W. Longcope and H. R. Strauss, Phys. Fluids B **5**, 2858 (1993).
[30]L. C. Lee and Z. F. Fu, J. Geophys. Res. **91**, 6807 (1986).
[31]E. R. Priest and T. G. Forbes, J. Geophys. Res. **97**, 16 757 (1992).
[32]J. B. Taylor, Phys. Rev. Lett. **35**, 1139 (1974).
[33]B. B. Kadomtsev, in *Plasma Physics and Controlled Fusion Research* (International Atomic Energy Agency, Vienna, 1977), Vol. I, p. 555.
[34]A. Bhattacharjee, R. L. Dewar, and D. A. Monticello, Phys. Rev. Lett. **45**, 347 (1980).
[35]M. N. Rosenbluth, R. Y. Dagazian, and P. H. Rutherford, Phys. Fluids **16**, 1894 (1973).
[36]See, for instance, P. K. Browning, Plasma Phys. Controlled Fusion **33**, 539 (1991), and other references therein.
[37]R. Rosner, W. H. Tucker, and G. S. Vaiana, Astrophys. J. **220**, 643 (1978).
[38]P. J. Cargill, Astrophys. J. **422**, 381 (1994).
[39]P. J. Cargill, in *Solar System Plasmas in Space and Time*, Geophysical Monograph 84, edited by J. L. Birch and J. H. Waite, Jr. (American Geophysical Union, Washington, DC, 1994), p. 21.
[40]See, for instance, A. G. Emslie, Sol. Phys. **121**, 105 (1989).
[41]R. P. Lin, R. A. Schwartz, S. R. Kane, R. M. Pelling, and K. C. Hurley, Astrophys. J. **283**, 421 (1984).
[42]X. Wang and A. Bhattacharjee, Phys. Rev. Lett. **70**, 1627 (1993).

GEOPHYSICAL RESEARCH LETTERS, VOL. 23, NO. 21, PAGES 2955-2958, OCTOBER 15, 1996

Role of photospheric footpoint shear in the impulsive dynamics of the solar corona

Z. W. Ma and A. Bhattacharjee

Department of Physics and Astronomy, University of Iowa, Iowa City

Abstract. A frequently observed feature of a solar flare is its impulsive growth from a relatively quiescent background. This imposes a significant constraint on magnetic reconnection models which propose to account for flares: not only should the time scale be fast, but it must also exhibit an impulsive character. It is shown that the spatial width of the shearing motion of the photospheric footpoints has a significant qualitative effect on the time-evolution of the current sheet and the reconnection electric field in a two-dimensional coronal loop with a neutral line. The reconnection dynamics for localized footpoint shear profiles does not show an impulsive phase and occurs on a Sweet-Parker time scale ($\sim \eta^{-1/2}$), whereas the reconnection dynamics with broad shear profiles occurs on a faster time scale with a weaker dependence on the resistivity ($\sim \eta^{-1/5}$), and furthermore, exhibits an impulsive phase due to flux pile-up near the reconnection layer.

Magnetic reconnection is frequently invoked as a possible mechanism for several observed features of coronal dynamics. It can cause a major rearrangement of magnetic topology and unleash a large reservoir of magnetic free energy that can be converted to thermal and bulk kinetic energy. The electric fields generated during reconnection can, in principle, accelerate particles to high energies. These attributes of magnetic reconnection are so appealing that we often invoke it as a panacea for several ill-understood problems of coronal dynamics.

Magnetic reconnection is often involved in the dynamical realization of intense concentrations of current density, called "current sheets." *Parker* [1972] proposed that such current sheets, which can grow from a smooth initial state (such as a vacuum magnetic field) due to shearing motions of the photospheric footpoints, can be an effective mechanism for coronal heating. Small events that produce current sheets and release energy ($\sim 10^{17}$ J) are classified as nanoflares. Current sheets are not only a defining feature of nanoflares, but are also expected to occur in the high-Lundquist-number corona during the development of microflares and flares if magnetic reconnection is the underlying mechanism liberating large amounts of energy ($\sim 10^{20}$ J for microflares and 10^{26} J for flares). Hence, the temporal development of current sheets (and the associated reconnection electric field) in a coronal loop has to be understood at a fundamental level before one can claim to have a correct time-dependent theory for flares.

Copyright 1996 by the American Geophysical Union.

Paper number 96GL02571.
0094-8534/96/96GL-02571$05.00

Observations provide some strong constraints on possible theoretical mechanisms for flares. (See, for instance, the monograph by *Tandberg-Hanssen and Emslie* [1988] and other references therein.) We emphasize two important *temporal signatures* that a theory of flares must reproduce. The first signature is a time scale fast enough to account for the rapidity of flaring, and the second, more subtle, feature is the time development of the process which typically (but not always) exhibits an *impulsive* character.

The main thrust of this paper is to present simulation studies of two-dimensional coronal loops in the high-Lundquist-number regime, elucidating the role that the shearing motion of the photospheric footpoints can play in realizing the two temporal signatures discussed above. We find that current sheets (and reconnection electric fields) that exhibit both of these features can, in principle, be obtained if the footpoints on the photosphere are driven by a sheared velocity profile of broad spatial extent that can cause flux pile-up [*Deluca and Craig,* 1992; *Wang et al.,* 1996] near the reconnection layer. However, if the sheared velocity profile is narrow and does not cause flux pile-up, the current sheet amplitude (and the reconnection electric field) in the nonlinear regime increases gradually (and algebraically) with time on a characteristic Sweet-Parker time scale ($\sim \eta^{-1/2}$), without showing a transition to a fast, impulsive phase.

To fix ideas, we consider a two-dimensional coronal loop with an X-type neutral line where current sheets can be realized by imposing footpoint displacements [*Low and Wolfson,* 1988; *Jensen,* 1989; *Karpen et al.,* 1990; *Finn and Lau,* 1991; *Vekstein, Priest, and Amari,* 1992; *Wang and Bhattacharjee,* 1994]. The temporal growth of current sheets in this geometry has been investigated in some depth by *Wang and Bhattacharjee* [1994] and *Ma et al.* [1995]. Figure 1 describes a coronal configuration in the $x-y$ plane, with the footpoints of the magnetic field intersecting the photosphere which is taken to be a plane perpendicular to the plane of the paper. For the class of footpoint displacements considered by Ma et al., the current density J_z grows as t^3, consistent with the analytical prediction in the high-Lundquist-number regime [*Wang and Bhattachar-*

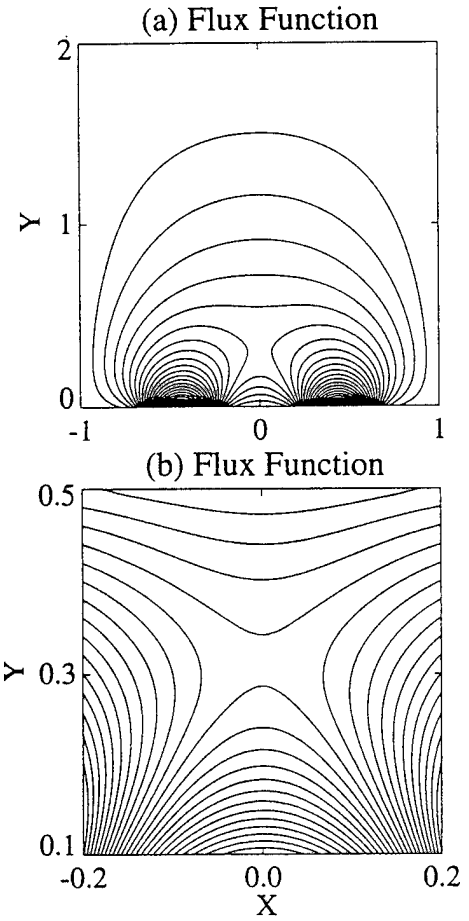

Figure 1. (a) Flux surfaces for a numerically generated force-free solution of the 2-D loop. (b) is a magnified picture of the region around the X-point in (a).

jee, 1994]. The reconnection dynamics in this simulation passes from a linear exponential phase to an algebraic phase in time, and this is too qualitatively gradual to account for the impulsive phase observed during a microflare or a flare.

In the present paper, we present numerical results from a two-dimensional, compressible, high-Lundquist-number simulation of the solar corona that go beyond the results reported by Ma et al. and bring the theory a step closer to the temporal signatures suggested by the observations. We assume the resistivity to be constant, without an ad hoc current-dependent or spatially localized enhancement factor. We integrate numerically the compressible resistive MHD equations using a Runge-Kutta finite-differencing scheme that has an accuracy of fourth-order both in time and space [*Ma and Bhattacharjee,* 1996]. Exploiting the symmetry of the initial conditions, we carry out the numerical simulation in half of the physical domain in the $x-y$ plane, with $x \in (0,1)a$ and $y \in (0,2)a$, where a is the characteristic system size. (A representative number for a is 50,000 km.) With a 181 × 221 array and a nonuniform grid, we are thus able to resolve $0.002a$ in x and $0.003a$ in y near the separatrix where the current sheet develops. The initial force-free equilibrium containing an X-type neutral line, shown in Figure 1, is obtained by solving Laplace's equation for a potential field in the whole physical domain as a boundary-value problem.

At the bottom boundary which represents the photosphere, we impose shearing motion of the footpoints according to the relation,

$$v_z(x, y=0, t) = v_0 \tanh\left(\frac{t}{t_0}\right) \exp\left[-\left(\frac{x-x_0}{\Delta x}\right)^2\right] , \quad (1)$$

where t_0 is a rapid "turn-on" time, $x = x_0$ is the point at which the outer separatrix intersects the x-axis in the simulation box (which is half of the physical domain), and Δx characterizes the spatial extent of the sheared profile. At the lateral boundary $x = 0$, we use symmetric as well as anti-symmetric boundary conditions for v_z. In the symmetric case, we impose footpoint velocities at the two points of intersection of the outer separatrix, either both out of the $(x-y)$ plane or both into the plane. (The inner separatrix is left unperturbed.) For the anti-symmetric case, the footpoint velocities at the two points of intersection of the outer separatrix are in opposite directions, with one footpoint moving out of the plane and the other moving into plane (also studied by Ma et al.) We find that the time-history for J_z is nearly the same for both types of boundary conditions. In the present paper, we only show the results for the symmetric case.

Localized Footpoint Shear and the Absence of an Impulsive Phase

Wang and Bhattacharjee [1994] show, by using a variant of the Sweet-Parker model, that in the nonlinear regime the current density J_z grows algebraically in time according to the relation,

$$J_z \cong J_0 \eta^{-1/2} (t/\tau_{sp})^3 , \quad (2)$$

where J_0 is a constant independent of η, $\tau_{sp} \equiv (\tau_A \tau_R)^{1/2}$ is the characteristic Sweet-Parker time scale, τ_A is the characteristic Alfvén time scale, and τ_R is the resistive diffusion time scale. In Figure 2, the dotted line indicates the time-evolution of the maximum amplitude of J_z at the separatrix in our simulation with $S \equiv \tau_R/\tau_A = 10^5$, and a localized footpoint shear profile with $\Delta x = 0.04a$. The dashed line represents the nonlinear analytical prediction given by equation (2).

We note that the growth of current sheets (or the re-

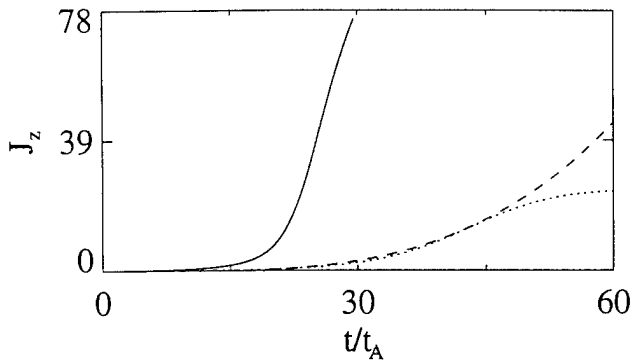

Figure 2. Maximum amplitudes (J_z) for localized shear profile (dotted line) and broad shear profile (solid line) as a function of time. The dashed line represents the analytically predicted t^3 behavior.

connection electric field) in this case is gradual throughout the nonlinear regime, and shows no impulsive phase. This gradual nature is also evident in the dotted line plot of Figure 3 in which the growth rate $d \ln J_z/dt$ is seen to decrease monotonically.

Broad Footpoint Shear and an Impulsive Phase

We now present numerical results with the same initial conditions as discussed above, except for the sheared footpoint velocity profile which is taken to be much broader, with $\Delta x = 0.16a$. ($\Delta x \sim 8000$ km.) The broader velocity profile implies simply that a larger bunch of field-lines are perturbed, and that the magnitude of the magnetic flux convected inwards from the boundary towards the separatrix is larger, causing flux pile-up.

Our recent analytical as well as numerical results [*Wang et al.*, 1996] for flux pile-up in the Harris sheet with inward boundary flows are relevant for the present simulations, and can be summarized as follows. The peak amplitude of the current sheet (and the reconnection electric field) in the Harris sheet exhibits a sudden transition from a sluggish linear growth phase to a rapid nonlinear phase in a characteristic time scale $\tau_N = (\tau_R \tau_0 \tau_A^3)^{1/5}$, where τ_0 is the characteristic time scale of the imposed boundary flow. It should be noted that the new characteristic time scale τ_N has a much weaker dependence on resistivity ($\sim \eta^{-1/5}$) than the characteristic time scale in the Sweet-Parker model ($\sim \eta^{-1/2}$).] Furthermore, the current sheet amplitude at the separatrix grows as $J_0(t/\tau_A)^2$ in the linear regime, and as $(\tau_R/\tau_0)^{1/2} J_0 (t/\tau_A)^{3/2}$ in the nonlinear regime, where J_0 is a constant, independent of η. The large multiplicative factor $(\tau_R/\tau_0)^{1/2}$ accounts for the suddenness in the enhancement of the current sheet amplitude in the nonlinear regime following a period of slower growth.

The loop geometry of the initial state in our present simulation makes a precise analytical treatment in the nonlinear regime difficult. However, the initial X-point structure stretches to form Y-points in the nonlinear regime, and resembles qualitatively the geometry realized in our Harris sheet studies. The time evolution of the maximum amplitude of the current density J_z (which is localized where $B_\perp [= (B_x^2 + B_y^2)^{1/2}]$ vanishes) is shown in Figure 2 (solid line). We find impulsive intensification of the current sheet amplitude at about $t = 20\tau_A$. (This impulsive intensification occurs for both J_z and $J_\perp$, but we only show the plot for J_z.) The impulsive enhancement in the time-evolution of the current density J_z (as well as the reconnection electric field $E_\parallel \equiv \mathbf{E} \cdot \mathbf{B}/B$ which is equal to E_z if $B_\perp = 0$) is evident by contrasting the dotted and solid line plots in Figure 3 for the growth rates $d \ln J_z/dt$ in the two cases of localized ($\Delta x = 0.04$) and broad ($\Delta x = 0.16$) footpoint shear, respectively.

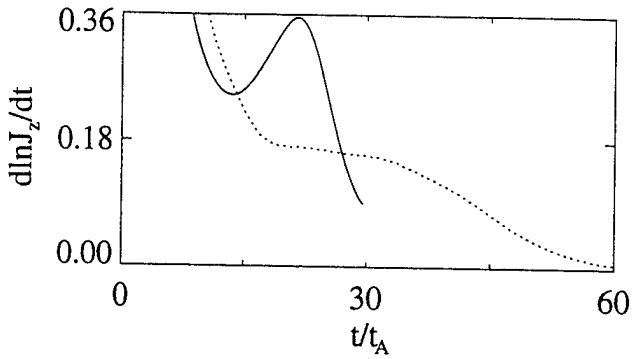

Figure 3. Time evolution of the growth rates of current sheets for localized shear profile (dotted line) and broad shear profile (solid line).

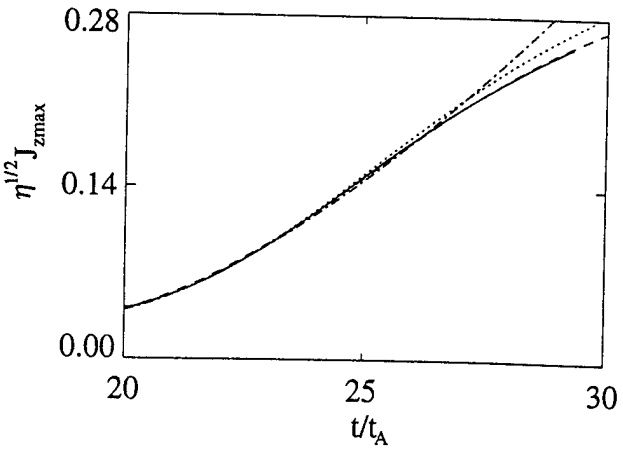

Figure 4. Time evolution of the scaled current $\eta^{1/2} J_{z\max}$ for $\eta = 10^{-5}$ (solid line), 2×10^{-5} (dotted line), and 4×10^{-5} (dashed line). The dashed-dotted line represents the analytical result.

In the case of broad shear, we conduct two additional runs with $\eta = 4 \times 10^{-5}$ and 2×10^{-5}. *Wang et al.* [1996] obtain the analytic expression for the nonlinear reconnected flux $\widetilde{\psi} \sim (t/\eta^{-1/5})^{5/2}$ which implies that the reconnection rate scales as $\eta^{-1/5}$, much faster than the Sweet-Parker rate. From this analytic expression, we obtain the nonlinear current sheet amplitude $J_z \sim \eta^{-1/2} t^{3/2}$. As shown in Figure 4, our numerical results on the dependence of the current sheet amplitude (as well as the reconnection rate) on resistivity and time are consistent with the analytical results of Wang et al. despite the differences in geometry.

In Figure 5 (a)–(c), we show plots of the current sheet amplitude J_z and the flux function for $S = 10^5$ at three instants of time ($t = 10, 20, 30\tau_A$). Note the evolution in the geometric features of the initial state, as the initial X-point structure is altered to form Y-points [*Syrovatsky*, 1971] which is a generic feature of reconnection in the high-Lundquist-number regime. (See Ma et al. for a further discussion of this point.) This geometric change is accompanied by a strong intensification of the current density which is distributed nonuniformly along the separatrix. (We draw the readers attention to the renormalization of the color scales in Figures 5(a)–(c).)

Even though we have not simulated realistic flare dynamics, the present results have significant qualitative implications for flare phenomena when magnetic reconnection is the underlying mechanism. We have demonstrated here that the nonlinear temporal dynamics of

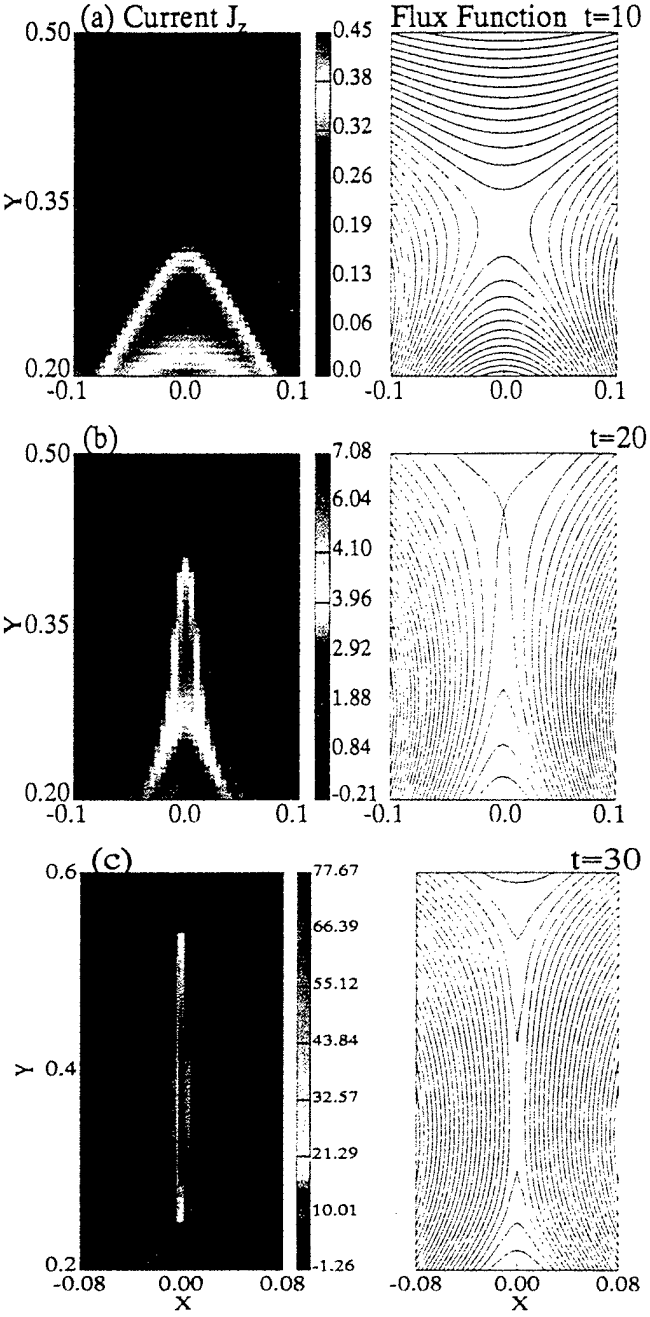

Figure 5. Image plots of the current and contour plots of flux function at $t = 10, 20$, and 30 for a broad shear profile with $S = 10^5$. The initial X-point geometry is altered to form Y-points as the current density is intensified. (Note the renormalization of the color scales in (a)–(c).)

current sheets (and the reconnection electric field) can be either gradual or impulsive depending on the spatial extent of the sheared velocity profiles of the footpoints. We remark that the impulsive enhancement of the current density is primarily due to the broad extent of the footpoint shear profile and not due to the plasma β; runs with the same boundary conditions but differing values of the plasma β (over the range $\sim 1-10\%$) show some quantitative but no qualitative differences. Collisionless effects in a generalized Ohm's law, discussed elsewhere [*Ma and Bhattacharjee*, 1996], can make the reconnection dynamics even more impulsive. There is substantial observational evidence that flares are strongly correlated with regions of pronounced velocity and magnetic field shear in the photosphere. (See, for instance, *Tandberg-Hanssen and Emslie* [1988], pp. 141–144 and other references therein.) We hope that the results of the present study will prompt a closer look at the data to determine the correlation between the time-development of flares and the spatial extent of the shear profiles.

Acknowledgments. This research is supported by NSF grant ATM 93-10157 and AFOSR grant F49620-93-1-0071. Supercomputer resources have been provided by the San Diego Supercomputing Center.

References

Deluca, E. E., and I. J. D. Craig, Magnetic reconnection in incompressible fluids, *Astrophys. J., 390,* 679, 1992.

Finn, J. M., and Y. T. Lau, Magnetohydrodynamic equilibria in vicinity of an X-type neutral line specified by footpoint shear, *Phys. Fluids B, 3,* 2675, 1991.

Jensen, T. H., Singularities of simple, force-free, line-tied magnetohydrodynamic equilibria, *Astrophys. J., 343,* 507, 1989.

Karpen, J. T., S. K. Antiochos, and C. R. Devore, On the formation of current sheets in the solar corona, *Astrophys. J., 356,* L67, 1990.

Low, B. C., and R. Wolfson, Spontaneous formation of electric current sheets and the origin of solar flares, *Astrophys. J., 324,* 574, 1988.

Ma, Z. W., C. S. Ng, X. Wang, and A. Bhattacharjee, Dynamics of current sheet formation and reconnection in two-dimensional coronal loops, *Phys. Plasmas, 2,* 3184, 1995.

Ma, Z. W. and A. Bhattacharjee, Fast impulsive reconnection and current sheet intensification due to the electron pressure gradients in a semi-collisional plasma, *Geophys. Res. Lett., 23,* 1673, 1996.

Parker, E. N., Topological dissipation and the small-scale fields in turbulent gases, *Astrophys. J., 174,* 499, 1972.

Syrovatsky, S. I., Formation of current sheets in a plasma with a frozen-in strong magnetic field, *Sov. Phys. JETP, 33,* 933, 1971.

Tandberg-Hanssen, E., and A. G. Emslie, *The Physics of Solar Flares,* Cambridge University Press, Cambridge, 1988.

Vekstein, G., E. Priest, and T. Amari, Formation of current sheets in force-free magnetic fields, *Astron. Astrophys., 243,* 492, 1991.

Wang, X., and A. Bhattacharjee, Current sheets and reconnection driven by footpoint motion in two-dimensional coronal loops with X-type neutral lines, *Astrophys. J., 420,* 415,1994.

Wang, X., Z. W. Ma, and A. Bhattacharjee, Fast reconnection and sudden enhancement of current sheets due to inward boundary flows, *Phys. Plasmas, 3,* 2129. 1996.

A. Bhattacharjee and Z. W. Ma, Department of Physics and Astronomy, University of Iowa, Iowa City, IA 52242. (e-mail: amitava@iowa.physics.uiowa.edu)

(received June 18, 1996; revised August 1, 1996; accepted August 12, 1996.)

THE ASTROPHYSICAL JOURNAL, 449:739–744, 1995 August 20
© 1995. The American Astronomical Society. All rights reserved. Printed in U.S.A.

SELF-CONSISTENCY CONSTRAINTS ON THE DYNAMO MECHANISM

A. BHATTACHARJEE AND Y. YUAN[1]
Department of Physics and Astronomy, University of Iowa, Iowa City, IA 52242
Received 1994 August 22; accepted 1995 March 7

ABSTRACT

It is shown that the requirements of self-consistency constrain the functional form of the turbulent dynamo field in an incompressible plasma. These requirements involve the back-reaction of the magnetic field through the Lorentz force in the momentum equation and the conservation laws of magnetohydrodynamic turbulence. The dynamo field is calculated in the weak-field limit when the turbulence is isotropic, as well as in the strong-field limit when the turbulence is anisotropic. For magnetic fields of a nontrivial topology, it is shown that the results of kinematic dynamo theory are strongly modified by the production of hyperresistivity (in the mean-field induction equation) which is left as a remnant after a near-cancellation between the alpha and beta effects. An interpolation formula for alpha quenching, encompassing weak-field and strong-field regimes, is proposed.

Subject headings: MHD — plasmas — turbulence

1. INTRODUCTION

The dynamo effect has been invoked as a mechanism for the generation and sustainment of astrophysical magnetic fields. The most well-developed branch of dynamo theory is the kinematic dynamo theory, which has been comprehensively discussed in several existing monographs (Moffatt 1978; Parker 1979; Krause & Rädler 1980; Zeldovich, Ruzmaikin, & Sokoloff 1983). Kinematic dynamo theory is essentially concerned with the question of growth (or decay) of a magnetic field $\boldsymbol{B}$, given a velocity field $\boldsymbol{v}$. The magnetic field is assumed to obey the induction equation of resistive magnetohydrodynamics (MHD),

$$\frac{\partial \boldsymbol{B}}{\partial t} = \nabla \times (\boldsymbol{v} \times \boldsymbol{B}) + \eta \nabla^2 \boldsymbol{B} , \tag{1}$$

where η is the resistivity of the conducting field. In turbulent systems, we can separate the variables $\boldsymbol{B}$ and $\boldsymbol{v}$ into an averaged part and a fluctuation, i.e.,

$$\boldsymbol{B} = \langle \boldsymbol{B} \rangle + \delta \boldsymbol{B} \equiv \boldsymbol{B}_0 + \delta \boldsymbol{B} , \tag{2a}$$

$$\boldsymbol{v} = \langle \boldsymbol{v} \rangle + \delta \boldsymbol{v} \equiv \boldsymbol{v}_0 + \delta \boldsymbol{v} , \tag{2b}$$

where $\langle \delta \boldsymbol{B} \rangle = \langle \delta \boldsymbol{v} \rangle = 0$. The angle brackets represent either an ensemble average or an average over the small-space and fast-time scales of the fluctuations. Averaging equation (1), we obtain

$$\frac{\partial \boldsymbol{B}_0}{\partial t} = \nabla \times (\boldsymbol{v}_0 \times \boldsymbol{B}_0) + \nabla \times \mathcal{E} + \eta \nabla^2 \boldsymbol{B}_0 , \tag{3}$$

where

$$\mathcal{E} \equiv \langle \delta \boldsymbol{v} \times \delta \boldsymbol{B} \rangle \tag{4}$$

is the turbulent dynamo field. Subtracting equation (3) from equation (1), we obtain an exact equation for the magnetic field fluctuation:

$$\frac{\partial}{\partial t} \delta \boldsymbol{B} = \nabla \times (\boldsymbol{v}_0 \times \delta \boldsymbol{B} + \delta \boldsymbol{v} \times \boldsymbol{B}_0) + \eta \nabla^2 \delta \boldsymbol{B} + [\nabla \times (\delta \boldsymbol{v} \times \delta \boldsymbol{B}) - \langle \delta \boldsymbol{v} \times \delta \boldsymbol{B} \rangle] . \tag{5}$$

In the "first-order smoothing" or "quasi-linear" approximation, the term in brackets on the right-hand side of equation (5) is neglected. The advantage of making this approximation is that it yields a linear equation for the fluctuation $\delta \boldsymbol{B}$, i.e.,

$$\frac{\partial}{\partial t} \delta \boldsymbol{B} = \nabla \times (\boldsymbol{v}_0 \times \delta \boldsymbol{B} + \delta \boldsymbol{v} \times \boldsymbol{B}_0) + \eta \nabla^2 \delta \boldsymbol{B} . \tag{6}$$

Equation (6) can be easily inverted to give $\delta \boldsymbol{B}$. If $\boldsymbol{v}_0$ is a constant, then it can be eliminated by a Galilean transformation. For turbulence that is isotropic but not reflectionally symmetric, one then obtains from equation (4) the well-known result (Moffatt 1978; Parker 1979; Krause & Rädler 1980; Zeldovich et al. 1983)

$$\mathcal{E} = \alpha_0 \boldsymbol{B}_0 - \beta_0 \boldsymbol{J}_0 , \tag{7}$$

where

$$\alpha_0 = -\frac{\tau}{3} \langle \delta \boldsymbol{v} \cdot \delta \boldsymbol{\omega} \rangle , \tag{8}$$

$$\beta_0 = \frac{\tau}{3} \langle |\delta \boldsymbol{v}|^2 \rangle , \tag{9}$$

and τ is an approximate eddy correlation time. Whereas the α_0-effect can amplify a seed magnetic field, the β_0-effect enhances the diffusion rate, typically to values much larger than that due to the classical resistivity η. The violation of mirror symmetry, a requirement for the α_0-effect, can occur due to the effects of natural cyclonic and nonuniform rotation on both small and large scales. If $\boldsymbol{v}_0$ is not spatially uniform, it cannot be eliminated by a Galilean transformation and leads to the so-called Ω-term in the mean-field induction equation (Moffatt 1978; Parker 1979; Zeldovich et al. 1983). If we write the mean magnetic field as the sum of a poloidal and a toroidal component, the Ω-term can generate a toroidal component from the poloidal component. The α_0-term, on the other hand,

[1] Permanent Address: Department of Applied Physics, Columbia University, New York, NY 10027.

regenerates the poloidal field from the toroidal component due to the effect of small-scale turbulence. In this paper, we focus on the α_0-effect, which is the heart of the mean-field amplification process.

From the inception of turbulent dynamo theory, it has been widely realized that the kinematic dynamo models are incomplete. If a weak seed magnetic field is amplified exponentially in time, the back-reaction of the magnetic field on the turbulent flow that generates it must eventually be taken into account and may alter qualitatively the predictions of kinematic dynamo theory. This realization gives rise to two important questions: first, what is the approximate magnitude of the mean magnetic field for which the kinematic theory ceases to be valid, and second, what happens to α_0 and β_0 when the kinematic theory no longer applies?

An answer to the first question might be that kinematic dynamo models will be invalid roughly when the energy of the mean magnetic field and the kinetic energy of the characteristic tubulent flow reach equipartition. It turns out that this answer is not supported by theory at the present time. This is because of the "Alfvén effect," a process by which the *small-scale* magnetic fluctuation energy reaches equipartition with the energy of the turbulent flow long before the large-scale magnetic field has picked up enough energy to reach equipartition with the turbulence (Pouquet, Frisch, & Leorat 1976). If one uses the simple estimate (Zeldovich 1957) (for which there is no rigorous justification in three dimensions),

$$\langle \delta B^2 \rangle^{1/2} \sim R_m^{1/2} B_0 , \tag{10}$$

where R_m is the magnetic Reynolds number, the Alfvén effect constrains the large-scale field B_0 by the inequality (Cattaneo & Vainshtein 1991; Vainshtein & Cattaneo 1992),

$$B_0 \lesssim \frac{\langle \delta U^2 \rangle^{1/2} \sqrt{4\pi\rho}}{R_m^{1/2}} . \tag{11}$$

Since R_m typically varies from 10^9 in stellar plasmas to 10^{14} in galactic plasmas, Cattaneo & Vainshtein claim that the inequality (11) severely restricts the magnitude of B_0 for which the predictions of kinematic dynamo models hold.

A different point of view, but one that also casts doubt on the validity of kinematic dynamo theory, is developed by Kulsrud & Anderson (1992). They note that in order for mean-field dynamo theory to be successful, it is important that the small-scale fluctuations be subdominant to the growing mean field. Unfortunately, they find the opposite to be the case in their calculations: there is much more energy on small scales than large, and the mean field generated by the kinematic dynamo effect is completely overwhelmed by the faster growing magnetic fluctuations. Thus, the "dynamo field quickly becomes unobservable under such conditions and the kinematic approximation fails before the mean field grows significantly" (Anderson & Kulsrud 1993, p. 1).

If we accept inequality (11) as the regime of validity of the kinematic theory, the next question is what happens to α_0 and β_0 when the kinematic approximation breaks down. In this paper, we attempt to answer this question by going beyond the kinetic approximation and imposing the constraints of self-consistency. We consider magnetic topologies that are more general than currently available models of α-quenching. (See, e.g., Rüdiger & Kichatinov 1993 and references therein, which assume that the mean magnetic field is uniform in space.) The standard result (7) is contained as a "weak-field" limit of our results. In the "strong-field" limit, we show that

$$\mathcal{E}_{\parallel} = \frac{\boldsymbol{B}_0}{B_0^2} \nabla \cdot \left(\kappa^2 \nabla \frac{\boldsymbol{J}_0 \cdot \boldsymbol{B}_0}{B_0^2} \right) , \tag{12}$$

where $\boldsymbol{J}_0 = \nabla \times \boldsymbol{B}_0$ and κ^2 is a positive definite functional. The form of equation (12) is popularly known as "hyperresistivity;" it conserves magnetic helicity and dissipates magnetic energy (Boozer 1986; Bhattacharjee & Hameiri 1986; Strauss 1986). We demonstrate here that for magnetic fields of nontrivial topology, hyperresistivity is left as a remnant after a remarkable near-cancellation between the α- and β-effects of kinematic dynamo theory.

The following is the plan of the paper. In § 2, we discuss the "weak-field" corrections to kinematic dynamo theory due to self-consistent dynamics. In § 3, we consider the "strong-field" limit in which the turbulence is strongly anisotropic. Though the results obtained in §§ 2 and 3 hold in different asymptotic regimes, we suggest that it is possible to interpolate between those regimes to obtain a form for $\mathcal{E}_{\parallel}$ that contains both asymptotic limits. In § 4, we give a derivation of the expression (12), using quasi-linear theory. We conclude in § 5 with a summary and a brief discussion of the implications of our results.

2. SELF-CONSISTENCY: THE WEAK-FIELD LIMIT

From the induction equation

$$\frac{\partial \boldsymbol{B}}{\partial t} + \nabla \times (\eta \boldsymbol{J} - \boldsymbol{v} \times \boldsymbol{B}) = 0 , \tag{13}$$

we can obtain the equation for the vector potential $\boldsymbol{A}(\boldsymbol{B} = \nabla \times \boldsymbol{A})$:

$$\frac{\partial \boldsymbol{A}}{\partial t} + \eta \boldsymbol{J} - \boldsymbol{v} \times \boldsymbol{B} = -\nabla \phi . \tag{14}$$

Here ϕ is a scalar function. Taking the scalar product of equation (13) with $\boldsymbol{A}$ and equation (14) with $\boldsymbol{B}$, and adding the results, we obtain

$$\frac{\partial}{\partial t} (\boldsymbol{A} \cdot \boldsymbol{B}) + 2\eta \boldsymbol{J} \cdot \boldsymbol{B} + \nabla \cdot \left(2\boldsymbol{E} \times \boldsymbol{A} + \frac{\partial \boldsymbol{A}}{\partial t} \times \boldsymbol{A} \right) = 0 , \tag{15}$$

where $\boldsymbol{E} = \eta \boldsymbol{J} - \boldsymbol{v} \times \boldsymbol{B}$. Averaging equation (15), we obtain

$$\frac{\partial}{\partial t} (\boldsymbol{A}_0 \cdot \boldsymbol{B}_0) + \frac{\partial}{\partial t} \langle \delta \boldsymbol{A} \cdot \delta \boldsymbol{B} \rangle + 2\eta \boldsymbol{J}_0 \cdot \boldsymbol{B}_0 + 2\eta \langle \delta \boldsymbol{J} \cdot \delta \boldsymbol{B} \rangle + \nabla \cdot \left(2\boldsymbol{E}_0 \times \boldsymbol{A}_0 + 2\langle \delta \boldsymbol{E} \times \delta \boldsymbol{A} \rangle + \frac{\partial \boldsymbol{A}_0}{\partial t} \times \boldsymbol{A}_0 + \left\langle \frac{\partial \delta \boldsymbol{A}}{\partial t} \times \delta \boldsymbol{A} \right\rangle \right) = 0 . \tag{16}$$

An alternative form for $\partial/\partial t(\boldsymbol{A}_0 \cdot \boldsymbol{B}_0)$ can be obtained by using the averaged form for equation (14), i.e.,

$$\frac{\partial \boldsymbol{A}_0}{\partial t} + \eta \boldsymbol{J}_0 - \boldsymbol{v}_0 \times \boldsymbol{B}_0 - \mathcal{E} = -\nabla \phi_0 , \tag{17}$$

and the averaged form for equation (13), i.e.,

$$\frac{\partial \boldsymbol{B}_0}{\partial t} + \nabla \times (\eta \boldsymbol{J}_0 - \boldsymbol{v}_0 \times \boldsymbol{B}_0 - \mathcal{E}) = 0 . \tag{18}$$

We take the scalar product of $\boldsymbol{B}_0$ with equation (17) and $\boldsymbol{A}_0$ with equation (18) and add the two equations. The result is

$$\frac{\partial}{\partial t}(\boldsymbol{A}_0 \cdot \boldsymbol{B}_0) + 2\eta \boldsymbol{J}_0 \cdot \boldsymbol{B}_0 - 2\boldsymbol{\mathscr{E}}_0 \cdot \boldsymbol{B}_0 + \nabla \cdot \left(2\boldsymbol{E}_0 \times \boldsymbol{A}_0 + \frac{\partial \boldsymbol{A}_0}{\partial t} \times \boldsymbol{A}_0\right) = 0 , \quad (19)$$

where $\boldsymbol{E}_0 = \eta \boldsymbol{J}_0 - \boldsymbol{v}_0 \times \boldsymbol{B}_0 - \boldsymbol{\mathscr{E}}$. Subtracting equation (19) from equation (16), we obtain

$$\boldsymbol{\mathscr{E}} \cdot \boldsymbol{B}_0 = -\eta \langle \delta \boldsymbol{J} \cdot \delta \boldsymbol{B} \rangle - \frac{1}{2}\frac{\partial}{\partial t}\langle \delta \boldsymbol{A} \cdot \delta \boldsymbol{B} \rangle - \nabla \cdot \left(\langle \delta \boldsymbol{E} \times \delta \boldsymbol{A} \rangle + \frac{1}{2}\left\langle \frac{\partial \delta \boldsymbol{A}}{\partial t} \times \delta \boldsymbol{A} \right\rangle \right), \quad (20)$$

which reduces to

$$\boldsymbol{\mathscr{E}} \cdot \boldsymbol{B}_0 = -\eta \langle \delta \boldsymbol{J} \cdot \delta \boldsymbol{B} \rangle + \langle \delta \boldsymbol{E} \cdot \delta \boldsymbol{B} \rangle . \quad (21)$$

Equation (20) was derived by Bhattacharjee & Hameiri (Bhattacharjee & Hameiri 1986; Hameiri & Bhattacharjee 1987) in their study of the dynamo effect in laboratory plasmas. Equations (20) and (21) are both exact relations and impose constraints on the allowable functional forms for $\boldsymbol{\mathscr{E}}$. (In deriving eq. [21], we have assumed, for simplicity, that the operations of spacetime differentiation and ensemble averaging commute.)

The back-reaction of the magnetic field on the turbulence that generates it modifies the kinematic relations (8) and (9). In the weak-field regime, when the turbulence can be regarded as isotropic, it has been shown (Pouquet et al. 1976; Vainshtein 1980; Zeldovich et al. 1983; Gruzinov & Diamond 1994) that the relation (7) changes to

$$\boldsymbol{\mathscr{E}} \simeq \alpha \boldsymbol{B}_0 - \beta \boldsymbol{J}_0 , \quad (22)$$

where

$$\alpha = -\frac{\tau}{3}\left(\langle \delta \boldsymbol{v} \cdot \delta \boldsymbol{\omega} \rangle - \frac{1}{\rho}\langle \delta \boldsymbol{J} \cdot \delta \boldsymbol{B} \rangle\right). \quad (23)$$

Vainshtein reports that $\beta = 2\beta_0$ for incompressible isotropic turbulence. However, if the effect of the perturbed pressure is taken into account, then it can be shown that

$$\beta = \beta_0 \quad (24)$$

(Avinash 1991; Gruzinov & Diamond 1994).

From equations (21) and (22), we obtain

$$\eta \langle \delta \boldsymbol{J} \cdot \delta \boldsymbol{B} \rangle = -\alpha B_0^2 + \beta \boldsymbol{J}_0 \cdot \boldsymbol{B}_0 + \langle \delta \boldsymbol{E} \cdot \delta \boldsymbol{B} \rangle . \quad (25)$$

Equation (25) should be compared with equation (9) of Gruzinov & Diamond (after a typographical error in eq. [9] of Gruzinov & Diamond is corrected). Gruzinov & Diamond maintain that their relation $\eta \langle \delta \boldsymbol{J} \cdot \delta \boldsymbol{B} \rangle = -\alpha B_0^2 + \beta \boldsymbol{J}_0 \cdot \boldsymbol{B}_0$ is an exact expression, but this is not so because it omits the last term in equation (25). It is this last term that yields hyperresistivity in the strong-field, anisotropic regime. Eliminating $\langle \delta \boldsymbol{J} \cdot \delta \boldsymbol{B} \rangle$ between equations (23) and (25), we obtain

$$\alpha = \frac{\alpha_0 + (\tau/3\rho\eta)(\beta \boldsymbol{J}_0 \cdot \boldsymbol{B}_0 + \langle \delta \boldsymbol{E} \cdot \delta \boldsymbol{B} \rangle)}{1 + (\tau/3\eta\rho)B_0^2} . \quad (26)$$

In the limit $(\tau/\eta\rho)B_0^2 \ll 1$, equation (26) gives $\alpha \simeq \alpha_0$, the kinematic result. (The second and third terms in the numerator of eq. [26] are much smaller than α_0 in this limit.)

We have not shown that equation (26) continues to hold in the strong-field regime, which is the subject of § 3. However, in anticipation of the results derived in §§ 3 and 4, we heuristically take the limit $(\tau/\eta\rho)B_0^2 \gg 1$ of equation (26). We obtain

$$\alpha \simeq \beta \frac{\boldsymbol{J}_0 \cdot \boldsymbol{B}_0}{B_0^2} + \frac{1}{B_0^2}\langle \delta \boldsymbol{E} \cdot \delta \boldsymbol{B} \rangle . \quad (27)$$

Substituting equation (27) in equation (22), we obtain

$$\boldsymbol{\mathscr{E}} = \left(\beta \frac{\boldsymbol{J}_0 \cdot \boldsymbol{B}_0}{B_0^2} + \frac{1}{B_0^2}\langle \delta \boldsymbol{E} \cdot \delta \boldsymbol{B} \rangle\right)\boldsymbol{B}_0 - \beta \boldsymbol{J}_0 , \quad (28)$$

$$= -\beta \boldsymbol{J}_{0\perp} + \frac{\boldsymbol{B}_0}{B_0^2}\langle \delta \boldsymbol{E} \cdot \delta \boldsymbol{B} \rangle . \quad (29)$$

The first term on the right-hand side of equation (29) is perpendicular to $\boldsymbol{B}_0$ and does not contribute to $\boldsymbol{\mathscr{E}}_\parallel$. Thus a significant cancellation has occurred between the β-effect (third term in eq. [28]) and (a part of) the α-effect (first term in eq. [28]), to yield

$$\boldsymbol{\mathscr{E}}_\parallel = \frac{\boldsymbol{B}_0}{B_0^2}\langle \delta \boldsymbol{E} \cdot \delta \boldsymbol{B} \rangle . \quad (30)$$

In §§ 3 and 4, we will demonstrate that equation (30) has the functional form of equation (12) in the strong-field limit.

3. SELF-CONSISTENCY: THE STRONG-FIELD LIMIT

As a large-scale magnetic field $\boldsymbol{B}_0$ grows in strength, the turbulent velocity field adjusts itself to the growing anisotropy induced by $\boldsymbol{B}_0$. To motivate the results that follow, we begin with a simple discussion of the main physical consequences of this growing anisotropy. For this purpose, neglecting collisional dissipation, we write the linearized equation for the fluctuating vector potential

$$\frac{\partial}{\partial t}\delta \boldsymbol{A} = \delta \boldsymbol{v} \times \boldsymbol{B}_0 - \nabla \delta \phi , \quad (31)$$

which can be resolved into parallel and perpendicular components:

$$\frac{\partial}{\partial t}\delta \boldsymbol{A}_\parallel = -\nabla_\parallel \delta \phi , \quad (32)$$

$$\frac{\partial}{\partial t}\delta \boldsymbol{A}_\perp = \delta \boldsymbol{v} \times \boldsymbol{B}_0 - \nabla_\perp \delta \phi . \quad (33)$$

Hence, we have

$$\boldsymbol{\mathscr{E}} = \left\langle \delta \boldsymbol{v} \times \nabla \times \int dt \frac{\partial \delta \boldsymbol{A}}{\partial t} \right\rangle , = \left\langle \delta \boldsymbol{v} \times \nabla \times \int dt (\delta \boldsymbol{v} \times \boldsymbol{B}_0 - \nabla_\perp \delta \phi) \right\rangle + \langle \delta \boldsymbol{v} \times \nabla \times (\delta \boldsymbol{A}_\parallel \hat{\boldsymbol{b}}) \rangle , \quad (34)$$

where $\hat{\boldsymbol{b}} \equiv \boldsymbol{B}_0/B_0$. The plasma beta, defined by $\beta_p \equiv 2p/B^2$, is a convenient parameter with which we can track the growth of the magnetic field (for fixed p). As discussed later in this section, for very large $\boldsymbol{B}_0$, the perturbed quantities $\delta \boldsymbol{A}_\perp$, $\delta \boldsymbol{v}_\parallel$, and $\delta \boldsymbol{J}_\perp$ are very small. Then equation (33) reduces to

$$\delta \boldsymbol{v} \times \boldsymbol{B}_0 \simeq \nabla_\perp \delta \phi . \quad (35)$$

When equation (35) holds, the first angle bracket on the right-hand side of equation (34) is nearly zero. Then the second term in equation (34) can be written, for incompressible plasmas, as

$$\mathscr{E}_{\parallel} = -\frac{\boldsymbol{B}_0}{B_0^2} \nabla \cdot \langle B_0 \, \delta A_{\parallel} \, \delta \boldsymbol{v}_{\perp} \rangle \, . \tag{36}$$

Equation (36) is a *local* relation, and the averaged terms under the divergence operator cannot be omitted in general by appealing to boundary conditions. We shall demonstrate that relation (36) (which is equivalent to the relation [30]) eventually leads to the functional form of equation (12).

Equation (36) can be derived in a more formally rigorous way by introducing a β_p-ordering on the self-consistent resistive MHD equations. For fusion plasmas, this procedure was first developed by Rosenbluth et al. (1976) and Strauss (1976), but a pedagogically more satisfactory derivation for astrophysical plasmas is given by Zank & Matthaeus (1993), who delineate three regimes $\beta_p \gg 1$, $\beta_p \sim 1$, and $\beta_p \ll 1$. We will use these regimes to classify the three cases of weak, moderate, and strong magnetic field, respectively. The case $\beta_p \gg 1$ is described well by fully three-dimensional, incompressible resistive MHD equations, and the turbulence is nearly isotropic. The results of § 2 pertain to this case. In the cases $\beta_p \sim 1$ and $\beta_p \ll 1$, the growth of $\boldsymbol{B}_0$ leads to the development of anisotropy. Assuming, for simplicity, that $\boldsymbol{B} \simeq \boldsymbol{B}_0 =$ constant to leading order, Zank & Matthaeus (1993) demonstrate formally that the effect of this large-scale field is "to introduce a preferred direction into the system that manifests itself by reducing the 'dimensionality' of the underlying incompressible description." Using a small parameter ϵ which is equal to the Alfvén Mach number $M_A = v_0/V_A \equiv v_0(\rho_0)^{1/2}/B_0$, Zank & Matthaeus derive, for $\beta_p \sim 1$, the system of equations

$$\nabla_{\perp} \cdot \boldsymbol{v}_{\perp} = 0 \, , \quad \nabla_{\perp} \cdot \boldsymbol{B}_{\perp} = 0 \, , \tag{37}$$

$$\left(\frac{\partial}{\partial t} + \boldsymbol{v}_{\perp} \cdot \nabla_{\perp} \right) \boldsymbol{v}_{\perp} = -\nabla_{\perp} \left(p + \frac{B_{\perp}^2}{2} \right) + (\boldsymbol{B}_{\perp} \cdot \nabla_{\perp}) \boldsymbol{B}_{\perp} \, , \tag{38}$$

$$\left(\frac{\partial}{\partial t} + \boldsymbol{v}_{\perp} \cdot \nabla_{\perp} \right) \boldsymbol{B}_{\perp} = (\boldsymbol{B}_{\perp} \cdot \nabla_{\perp}) \boldsymbol{v}_{\perp} + \eta \nabla^2 \boldsymbol{B}_{\perp} \, , \tag{39}$$

where $\boldsymbol{B} = \boldsymbol{B}_{\perp} + \boldsymbol{B}_0 \hat{z}$ and $\boldsymbol{v} = \boldsymbol{v}_{\perp}$ to leading order. In this ordering, $\boldsymbol{A}_{\perp}$ and $v_{\parallel}$ are zero to leading order, as asserted earlier. It is then possible to write $\boldsymbol{v} = \nabla_{\perp} \hat{\phi}(x, y, t) \times \hat{z}$, where $\hat{\phi} = -\phi/B_0$ and $\boldsymbol{B}_{\perp} = \nabla_{\perp} A_{\parallel}(x, y, t) \times \hat{\boldsymbol{b}}$. Hence, we have

$$\mathscr{E}_{\parallel} = \langle \delta \boldsymbol{v} \times \delta \boldsymbol{B} \rangle_{\parallel} = \frac{\boldsymbol{B}_0}{B_0^2} \nabla \cdot \langle -B_0 \, \delta \boldsymbol{v}_{\perp} \, \delta A_{\parallel} \rangle$$

$$= \frac{\boldsymbol{B}_0}{B_0^2} \nabla \cdot \langle B_0 \, \delta \hat{\phi} \, \delta \boldsymbol{B}_{\perp} \rangle \, , \tag{40}$$

which is identical to equation (36).

In the next section, we will use quasi-linear theory to show that equation (40) has the form of equation (12), with κ^2 related to the spectrum of fluctuations.

4. DERIVATION OF HYPERRESISTIVITY FROM QUASI-LINEAR THEORY

We introduce Elsasser variables $\boldsymbol{Z}^+ = \boldsymbol{B}_{\perp}/\rho_0^{1/2} + \boldsymbol{v}_{\perp}$ and $\boldsymbol{Z}^- = \boldsymbol{B}_{\perp}/\rho_0^{1/2} - \boldsymbol{v}_{\perp}$. Then, neglecting collisional dissipation equations (37)–(39) can be combined to give

$$\nabla_{\perp} \cdot \boldsymbol{Z}^+ = 0 \, , \quad \nabla_{\perp} \cdot \boldsymbol{Z}^- = 0 \, , \tag{41}$$

$$\frac{\partial}{\partial t} \boldsymbol{Z}^+ - \boldsymbol{Z}^- \cdot \nabla_{\perp} \boldsymbol{Z}^+ = -\nabla_{\perp} P_* \, , \tag{42}$$

and

$$\frac{\partial}{\partial t} \boldsymbol{Z}^- + \boldsymbol{Z}^+ \cdot \nabla_{\perp} \boldsymbol{Z}^- = \nabla_{\perp} P_* \, , \tag{43}$$

where $P_* \equiv p/\rho_0 + (\boldsymbol{Z}^+ + \boldsymbol{Z}^-)^2/8$. Hereafter, we set $\rho_0 = 1$. We assume that $\langle \boldsymbol{B} \rangle = B_y(x)\hat{y} + B_0 \hat{z}$ and that the inhomogeneity depends only on x. Since $\delta \boldsymbol{B}_{\perp} = \nabla \delta\psi \times \hat{\boldsymbol{b}}$ (where $A_{\parallel} \equiv \psi$) and $\delta \boldsymbol{v}_{\perp} = \nabla \delta\hat{\phi} \times \hat{\boldsymbol{b}}$, we have

$$\delta \boldsymbol{Z}^+ = \nabla \xi_+ \times \hat{\boldsymbol{b}} \, , \quad \delta \boldsymbol{Z}^- = \nabla \xi_- \times \hat{\boldsymbol{b}} \, , \tag{44}$$

where $\xi_+ = \delta\psi + \delta\hat{\phi}$ and $\xi_- = \delta\psi - \delta\hat{\phi}$. Then the turbulent electromotive force $\mathscr{E}$ can be written as

$$\mathscr{E} = \langle \delta \boldsymbol{v}_{\perp} \times \delta \boldsymbol{B}_{\perp} \rangle = \frac{\hat{\boldsymbol{b}}}{4} \nabla \cdot \langle \xi_+ \, \delta \boldsymbol{Z}^- - \xi_- \, \delta \boldsymbol{Z}^+ \rangle \, . \tag{45}$$

We Fourier-analyze along y- and z-directions, and write

$$\xi_+ \, \delta \boldsymbol{Z}^- = \frac{1}{(2\pi)^2} \sum_{k_y, k_y', k_z, k_z'} \xi_{+ k_y k_z}(x, t) \delta \boldsymbol{Z}^-_{k_y' k_z'}(x, t) \times \exp \left[i(k_y + k_y')y - i(k_z + k_z')z \right]$$

$$\simeq \frac{1}{(2\pi)^2} \sum_{\boldsymbol{k}} \xi_{+ k_y k_z}(x, t) \delta \boldsymbol{Z}^{-*}_{k_y k_z}(x, t) \, , \tag{46}$$

where * denotes complex conjugation. Similarly, we have

$$\xi_- \, \delta \boldsymbol{Z}^+ \simeq \frac{1}{(2\pi)^2} \sum_{\boldsymbol{k}} \xi_{- k_y k_z}(x, t) \delta \boldsymbol{Z}^{+*}_{k_y k_z}(x, t) \, , \tag{47}$$

where $\boldsymbol{k} = (k_y, k_z)$. From equation (44), we obtain

$$\nabla_{\perp}^2 \xi_+ = \hat{\boldsymbol{b}} \cdot \nabla \times \delta \boldsymbol{Z}^+ \, , \tag{48}$$

which gives

$$\xi_{+ \boldsymbol{k}} = k_{\perp}^{-2} \hat{\boldsymbol{b}} \cdot [i\boldsymbol{k} \times \delta \boldsymbol{Z}^+_{\boldsymbol{k}}(x, t) + \nabla \times \delta \boldsymbol{Z}^+_{\boldsymbol{k}}(x, t)] \, . \tag{49}$$

Similarly, from

$$\nabla_{\perp}^2 \xi_- = \hat{\boldsymbol{b}} \cdot \nabla \times \delta \boldsymbol{Z}^- \, , \tag{50}$$

we obtain

$$\xi_{- \boldsymbol{k}} = k_{\perp}^{-2} \hat{\boldsymbol{b}} \cdot [i\boldsymbol{k} \times \delta \boldsymbol{Z}^-_{\boldsymbol{k}}(x, t) + \nabla \times \delta \boldsymbol{Z}^-_{\boldsymbol{k}}(x, t)] \, . \tag{51}$$

It is clear by inspection of equations (45), (48), and (51) that in order to compute $\mathscr{E}$, we need a renormalized turbulence theory to deal with the nonlinear terms $\boldsymbol{Z}^- \cdot \nabla \boldsymbol{Z}^+$, $\boldsymbol{Z}^+ \cdot \nabla \boldsymbol{Z}^-$, and $(\boldsymbol{Z}^+ + \boldsymbol{Z}^-)^2$ in equations (42)–(43). For simplicity, in order to keep the underlying physics as transparent as possible, we choose to use the quasi-linear approximation, neglecting mode-coupling effects. The linear propagators in our derivation can be shown to be broadened by nonlinear effects (Strauss 1986; Tetreault 1989). In the quasi-linear approximation, equations (42) and (43) yield

$$\left(\frac{\partial}{\partial t} - \langle \boldsymbol{Z}^- \rangle \cdot \nabla \right) \delta \boldsymbol{Z}^+ = \delta \boldsymbol{Z}^- \cdot \nabla_{\perp} \langle \boldsymbol{Z}^+ \rangle \, , \tag{52}$$

$$\left(\frac{\partial}{\partial t} + \langle \boldsymbol{Z}^+ \rangle \cdot \nabla \right) \delta \boldsymbol{Z}^- = -\delta \boldsymbol{Z}^+ \cdot \nabla_{\perp} \langle \boldsymbol{Z}^- \rangle \, , \tag{53}$$

where we have used the condition $\delta p + \langle \boldsymbol{B} \rangle \cdot \delta \boldsymbol{B} = 0$, valid for a wide class of resistive modes (see, e.g., Hameiri & Bhatta-

charjee 1987). In Fourier space, equations (52) and (53) give

$$\left(\frac{\partial}{\partial t} - ik_y\langle Z^-\rangle_y\right)\delta Z_k^+ = \delta Z_k^- \cdot \nabla_\perp \langle Z^+\rangle \tag{54}$$

and

$$\left(\frac{\partial}{\partial t} + ik_y\langle Z^+\rangle_y\right)\delta Z_k^- = -\delta Z_k^+ \cdot \nabla_\perp \langle Z^-\rangle\,, \tag{55}$$

respectively. Equations (54) and (55) can then be easily inverted. Formally, we write

$$\delta Z_k^+ = G_k^- \delta Z_k^- \cdot \nabla \langle Z^+\rangle \tag{56}$$

and

$$\delta Z_k^- = -G_k^+ \delta Z_k^+ \cdot \nabla \langle Z^-\rangle\,, \tag{57}$$

where $G_k^\pm$ are the Green's functions, given by

$$G_k^\pm = (\gamma_k \pm ik_y\langle Z^\pm\rangle_y)^{-1}\,. \tag{58}$$

Using equations (56) and (57) in equations (49) and (51), respectively, we obtain

$$\xi_{+k} \simeq k_\perp^{-2} G_k^- \delta Z_k^- \cdot \nabla \langle J_\parallel\rangle \tag{59}$$

and

$$\xi_{-k} \simeq -k_\perp^{-2} G_k^+ \delta Z_k^+ \cdot \nabla \langle J_\parallel\rangle\,. \tag{60}$$

Substituting equations (59) and (60) in equation (45), we obtain

$$\mathscr{E} \simeq \frac{\boldsymbol{B}_0}{B_0^2} \nabla \cdot \mathbf{K} \cdot \nabla\left(\frac{\boldsymbol{J}_0 \cdot \boldsymbol{B}_0}{B_0^2}\right), \tag{61}$$

where

$$\mathbf{K} = \frac{B_0^2}{(4\pi)^2} \sum_k k_\perp^{-2}(G_k^- \delta Z_k^{-*} \delta Z_k^- + G_k^+ \delta Z_k^{+*} \delta Z_k^+)\,. \tag{62}$$

An equation similar to equation (62) has been derived earlier by Tetreault (1989) in the context of MHD clump turbulence in toroidal plasmas, confined by a strong toroidal field $\boldsymbol{B}_0$. If the fluctuations are dominated by one component orthogonal to $\boldsymbol{B}_0$ (say, the x-component) which is a case of considerable physical interest, then equation (61) can be approximated by the functional form of equation (12). We note that equation (12) obeys the integral relations,

$$\int_V \mathscr{E} \cdot \boldsymbol{B}_0\, d\boldsymbol{r} = \int_S \kappa^2 \nabla\left(\frac{\boldsymbol{J}_0 \cdot \boldsymbol{B}_0}{B_0^2}\right) \cdot d\boldsymbol{a}\,, \tag{63}$$

and

$$\int_V \mathscr{E} \cdot \boldsymbol{J}_0\, d\boldsymbol{r} = -\int_V \kappa^2 \left[\nabla\left(\frac{\boldsymbol{J}_0 \cdot \boldsymbol{B}_0}{B_0^2}\right)\right]^2 d\boldsymbol{r} + \int_S \kappa^2 \left(\frac{\boldsymbol{J}_0 \cdot \boldsymbol{B}_0}{B_0^2}\right) \nabla\left(\frac{\boldsymbol{J}_0 \cdot \boldsymbol{B}_0}{B_0^2}\right) d\boldsymbol{r}\,, \tag{64}$$

where S denotes the surface of V. The physical significance of equations (63) and (64) is clear: the turbulent dynamo field $\mathscr{E}$ neither creates nor destroys helicity in any volume V but dissipates energy (Boozer 1986, 1993; Bhattacharjee & Hameiri 1986) within the volume. (If the dynamo field would have created or destroyed helicity, there would have been a volume term in eq. [63]. There is a volume term in eq. [64], and it is negative definite, representing dissipation.)

We discuss the physical implications of the calculation given above in the context of astrophysical plasmas. In the presence of a large-scale background magnetic field $\boldsymbol{B}_0$, the plasma turbulence is envisioned to be a bath of Alfvénic fluctuations. Indeed, if we set the right-hand side of equations (52) and (53) to zero, we obtain uncoupled Alfvén wave fluctuations which obey the relation $\delta \boldsymbol{v} = \pm \delta \boldsymbol{B}$. Observations of incompressible MHD turbulence in the solar wind (see, for instance, Belcher & Davis 1971; Burlaga & Turner 1979) indicate a tendency of alignment or anti-alignment between the fluctuations $\delta \boldsymbol{v}$ and $\delta \boldsymbol{B}$. Dobrowolny, Mangeney, & Veltri (1980) have shown from considerations of the inertial range of the turbulence (ignoring source or dissipation terms) that this tendency is a general consequence of the dynamical relaxation of self-consistent MHD turbulence if the initial excitation favors one type of Alfvén fluctuation (+ or −). It should be emphasized that alignment (or anti-alignment) is merely a tendency and not realized in practice. (If this asymptotic state were realizable, then according to eq. [23], α would be exactly zero.) However, the tendency in itself is indicative of the fact that $\mathscr{E}$ calculated by the kinematic theory will be strongly reduced by the Alfvén effect. In the neighborhood of the asymptotic state, one expects nonlinear mode-coupling effects to be weak (Dobrowolny et al. 1980). There are many Fourier modes in the turbulent bath, and quasi-linear theory, which sums over the modes but neglects nonlinear mode-coupling effects, is a reasonable first approximation. In the context of this physical picture, our calculation shows that when the two types of propagating Alfvén wave fluctuations are coupled by the terms on the right-hand side of equations (52) and (53), we obtain the result (61) in which hyperresistivity is left as a remnant after a near-exact cancellation between the alpha and beta effects of the kinematic theory.

Before we conclude this section, we draw the attention of the reader to an instructive discussion of the Alfvén effect in Biskamp's recent monograph (Biskamp 1993). Biskamp gives a qualitative discussion of the importance of hyperresistivity in the context of MHD turbulence and the inverse cascade phenomenon that underlies the conservation of magnetic helicity.

5. CONCLUSIONS

In this paper, we have examined the constraints imposed by self-consistency on the turbulent dynamo in the weak-field as well as strong-field regimes. Synthesizing the results of §§ 2–4, we propose the interpolation formula

$$\alpha = \left\{\alpha_0 + \frac{\tau}{3\rho\eta}\left[\beta \boldsymbol{J}_0 \cdot \boldsymbol{B}_0 + \nabla \cdot \kappa^2 \nabla\left(\frac{\boldsymbol{J}_0 \cdot \boldsymbol{B}_0}{B_0^2}\right)\right]\right\} \times \left[1 + \frac{\tau}{3\rho\eta} B_0^2\right]^{-1}. \tag{65}$$

As demonstrated in §§ 2 and 3, when equation (65) is substituted in equation (22), there is a near-cancellation between the α- and β-effects, and in the strong-field limit, we are left with the functional form of equation (12) for $\mathscr{E}_\parallel$, known as hyperresistivity. Even though our derivation of hyperresistivity is based on the quasi-linear approximation, we believe that the functional dependencies of this term on mean physical variables is robust because it is consistent with well-known properties of three-dimensional MHD turbulence (Taylor 1974; Pouquet et al. 1976; Matthaeus & Montgomery 1980; Boozer 1986; Bhattacharjee & Hameiri 1986). Hyperresistivity does

not amplify either magnetic flux or energy. It can, for example, convert toroidal flux to poloidal flux as long as the conversion is consistent with helicity conservation and dissipates magnetic energy in the process.

While these conclusions pose critical challenges for traditional turbulent MHD dynamo models, they do not negate the relevance of the traditional theory for all astrophysical magnetic fields. The mechanism and affectiveness of the saturation mechanism discussed in this paper may not apply in all circumstances, particularly if the mean magnetic field is very weak. Thus, the galactic dynamo problem may require considerations rather different from those relevant to the solar or planetary dynamo. Field (1994) has recently given a useful summary of the issues raised by recent criticisms of the dynamo theory for galactic magnetic fields, and it is clear that a final resolution of the problem of origin of galactic magnetic fields depends on reliable calculations of alpha and beta over relevant timescales, after freeing "the classical theory from having to assume the first-order smoothing approximation." The present self-consistent calculation is a step in that direction.

This research is supported by the Air Force Office of Scientific Research grant F49620-93-1-0071 and the National Science Foundation grant ATM-9310157. We thank F. Cattaneo for stimulating our interest in this problem and helpful discussions, S. Spangler for useful pointers to the literature. We also thank an anonymous referee for constructive criticism that helped improve the presentation of this paper.

REFERENCES

Anderson, S. W., & Kulsrud, R. M. 1993, Solar and Planetary Dynamos (Cambridge: Cambridge Univ. Press)
Avinash, K. 1991, Phys. Fluids, B3, 2150
Belcher, J. W., & Davis, L., Jr. 1971, J. Geophys. Res.,76, 3534
Bhattacharjee, A., & Hameiri, E. 1986, Phys. Rev. Lett., 57, 206
Biskamp, D. 1993, Nonlinear Magnetohydrodynamics (Cambridge: Cambridge Univ. Press)
Boozer, A. H. 1986, J. Plasma Phys., 35, 133
———. 1993, Phys. Fluids B, 5, 2271
Burlaga, L. F., & Turner, J. M. 1979, J. Geophys. Res., 81, 73
Cattaneo, F., & Vainshtein, S. I. 1991, ApJ, 376, L21
Dobrowolny, M., Mangeney, A., & Veltri, P. 1980, Phys. Rev. Lett., 45, 144
Field, G. B. 1994, in Proc. 1994 Internat. Conf. on Plasma Physics, in press
Gruzinov, A. V., & Diamond, P. H. 1994, Phys. Rev. Lett., 72, 1651
Hameiri, E., & Bhattacharjee, A. 1987, Phys. Fluids, 30, 1743
Krause, F., & Rädler, K.-H. 1980, Mean-Field Magnetohydrodynamics and Dynamo Theory (Oxford: Pergamon)
Kulsrud, R. M., & Anderson, S. W. 1992, ApJ, 396, 606
Matthaeus, W. H., & Montgomery, D. C. 1980, Ann. N.Y. Acad. Sci., 357, 203
Moffatt, H. K. 1978, Field Generation in Electric Conducting Fluids (Cambridge: Cambridge Univ. Press)
Parker, E. N. 1979, Cosmic Magnetic Fields, Their Origin and Activity (Oxford: Clarendon)
Pouquet, A., Frisch, U., & Leorat, J. 1976, J. Fluid Mech., 77, 321
Rosenbluth, M. N., Monticello, D. A., Strauss, H. R., & White, R. B. 1976, Phys. Fluids, 19, 1987
Rüdiger, G., & Kichatinov, L. L. 1993, A&A, 269, 581
Strauss, H. R. 1976, Phys. Fluids, 19, 134
———. 1986, Phys. Fluids, 29, 3668
Taylor, J. B. 1974, Phys. Rev. Lett., 33, 1139
Tetreault, D. 1989, Phys. Fluids B, 1, 511
Vainshtein, S. I. 1980, Magnetohydrodynamics, 16, 111
Vainshtein, S. I., & Cattaneo, F. 1992, ApJ, 393, 165
Zank, G. P., & Matthaeus, W. H. 1993, Phys. Fluids, A5, 257
Zeldovich, Ya. B. 1957, Sov. Phys.–JETP, 4, 460
Zeldovich, Ya. B., Ruzmaikin, A. A., & Sokoloff, D. D. 1983, Magnetic Fields in Astrophysics (New York: Gordon & Breach)

THE ASTROPHYSICAL JOURNAL, 465:845–854, 1996 July 10
© 1996. The American Astronomical Society. All rights reserved. Printed in U.S.A.

INTERACTION OF SHEAR-ALFVÉN WAVE PACKETS: IMPLICATION FOR WEAK MAGNETOHYDRODYNAMIC TURBULENCE IN ASTROPHYSICAL PLASMAS

C. S. NG AND A. BHATTACHARJEE
Department of Physics and Astronomy, University of Iowa, Iowa City, IA 52242
Received 1995 August 24; accepted 1996 January 24

ABSTRACT

The nonlinear interaction of shear-Alfvén wave packets is a fundamental physical process underlying incompressible magnetohydrodynamics turbulence, as emphasized in the Iroshnikov-Kraichnan theory. In the weak turbulence limit, we give a detailed analytical and numerical treatment of the interaction between two colliding shear-Alfvén wave packets in the presence of a strong and uniform magnetic field $\boldsymbol{B} = B_0 \hat{z}$. Using the ideal MHD equations, it is shown that three-wave interactions are generally nonzero if the $k_z = 0$ Fourier components of the wave packets are nonzero. From the reduced MHD equations, we calculate in closed form the three-wave and four-wave interaction terms, and show the latter to be generally asymptotically subdominant if the wave packets have no $k_z = 0$ component. Our results on the generic dominance of three-wave interactions contradict recent claims by Sridhar & Goldreich (1994) who have argued that three-wave interactions are empty and that the Iroshnikov-Kraichnan theory is incorrect because it describes weak three-wave turbulence. The principal implication of our results is that the Iroshnikov-Kraichnan theory is still a suitable point of departure for the study of Alfvénic turbulence in the interstellar medium.

Subject headings: MHD — plasmas — turbulence — waves

1. INTRODUCTION

Incompressible magnetohydrodynamics (MHD) is the standard model for theoretical studies of astrophysical plasma turbulence. Although the applicability of the model to the high-beta compressible plasmas commonly encountered in the interstellar medium (ISM) or the solar wind is open to serious questions, the predictions of this model must be understood before one can attempt a more complete dynamical theory that includes compressibility and kinetic effects. An issue that has crucial implications for theory as well as observations is the form of the turbulent energy spectrum. Although the incompressible MHD equations reduce to the incompressible Navier-Stokes equation in the limit of zero magnetic field, it cannot be claimed that the energy spectrum (in the absence of intermittency corrections) must be Kolmogorov even if the magnetic field is weak. If the magnetic field is nonzero, Alfvén waves play an essential role in mediating the dynamics of turbulent fluctuations. In the Iroshnikov-Kraichnan (hereafter, IK) theory (Iroshnikov 1963; Kraichnan 1965), the collision of oppositely directed Alfvén wave packets provides a mechanism for the cascade of energy to short wavelengths. Scaling arguments (reviewed below) then yield the IK energy spectrum $E(k) \propto k^{-3/2}$. Numerical simulations of MHD turbulence have provided some evidence in support of the IK spectrum (see, for instance, Biskamp 1993), but the evidence is not yet conclusive.

With a view to developing a theoretical basis for turbulence in the ISM (see, for example, reviews by Rickett 1990 and Narayan 1992), Sridhar & Goldreich (1994) (hereafter SG) have revisited the IK theory, and claim that the theory is incorrect. There are three main steps in SG's refutation of the IK theory. SG first note that in the weak-turbulence limit, the main property needed to derive the IK spectrum is that the interaction between two oppositely propagating wave packets (with amplitudes of the same order) produces a distortion of the wave packets that is proportional to the square of the amplitude. Second, they note that the distortion that scales as the square of the amplitude must be attributed to resonant three-wave interactions. Finally, they make the crucial (and surprising) claim that resonant three-wave interactions are absent in the theory because the resonant coupling coefficients vanish. Whereas the first two steps are implicit in the IK theory, it is in the third step that SG differ sharply with IK. Having concluded that three-wave interactions are absent, SG proceed to consider four-wave interactions which yield distortions in colliding wave packets proportional to the *cube* of the amplitude of the wave packets. This cubic dependence of the distortion of the amplitude changes the energy spectrum to $E(k) \propto k^{-4/3}$, which should be contrasted with the IK spectrum $E(k) \propto k^{-3/2}$. In a subsequent paper, Goldreich & Sridhar (1995) have built further on the weak-turbulence theory of SG to develop a strong turbulence theory for the ISM.

Much of the weight of SG's critique of the IK theory rests on their crucial physical argument regarding the absence of resonant three-wave interactions. Montgomery & Matthaeus (1995) have recently expressed disagreement with SG by qualitative arguments that strongly suggest that the contribution from three-wave resonant interactions do not vanish. Though the physical arguments of Montgomery & Matthaeus have considerable merit, it seems to us that a definite resolution of this controversy may require more than qualitative arguments on the crucial question of whether three-wave or four-wave interactions dominate the nonlinear interaction of two colliding Alfvén wave packets. In this paper, we investigate in detail the three- and four-wave processes that mediate the collision of two wave packets propagating in opposite directions. We calculate analytically the three- and four-wave interactions in closed form using pertubation theory, and verify the analytical calculations numerically. Our principal conclusion is that resonant three-wave interactions do not generally vanish. The IK theory remains, in our view, a useful point of departure for the study of astrophysical plasma turbulence.

The following is a plan of this paper. In § 2, following SG, we review the heuristic scaling that governs three-wave and four-wave interactions and their implications for the energy spectrum. In § 3, we consider the collision of two oppositely directed Alfvén wave packets according to the ideal MHD equations. We solve for the distortion of the wave packets by perturbation theory and show that the first-order distortion, which is attributable to three wave interactions, is nonzero in general. For simplicity, we consider the reduced MHD (hereafter, RMHD) equations which can be deduced rigorously from

the full MHD equations in the presence of a large and uniform magnetic field, $B_0 \hat{z}$, assuming that the plasma β is less than unity (Rosenbluth et al. 1976; Strauss 1976; Zank & Matthaeus 1993). The RMHD equations can describe the nonlinear dynamics of shear-Alfvén waves in a low-β plasma. Shebalin, Matthaeus, & Montgomery (1983) have shown numerically that in the presence of a strong magnetic field, an initially isotropic spectrum tends to become strongly anisotropic. Thus, the evolving turbulent MHD dynamics provides a natural separation of spatial scales (with slow variation parallel and rapid variation perpendicular to the strong magnetic field) that supports the RMHD ordering. The advantage of the RMHD equations is that they can be solved analytically for the first-order (three-wave) as well the second-order (four-wave) interaction terms in closed form (§ 4). In the weak-turbulence limit, these analytical results demonstrate that the distortion of colliding wave packets due to four-wave interactions is much smaller than that due to three-wave interactions. In § 5, the analytical results from § 4 are compared and shown to be in good agreement with numerical RMHD simulations. We conclude with a summary and a discussion of the implications of our result in § 6.

2. HEURISTICS OF THREE- AND FOUR-WAVE INTERACTIONS: IMPLICATIONS FOR THE ENERGY SPECTRUM

Following SG, we begin with a heuristic discussion of the implications of three- and four-wave interactions for the energy spectrum of incompressible MHD turbulence governed by the equations,

$$\frac{\partial \boldsymbol{v}}{\partial t} + \boldsymbol{v} \cdot \nabla \boldsymbol{v} = -\nabla p + \boldsymbol{B} \cdot \nabla \boldsymbol{B} + \nu \nabla^2 \boldsymbol{v} , \qquad \frac{\partial \boldsymbol{B}}{\partial t} + \boldsymbol{v} \cdot \nabla \boldsymbol{B} = \boldsymbol{B} \cdot \nabla \boldsymbol{v} + \eta \nabla^2 \boldsymbol{B} , \qquad \nabla \cdot \boldsymbol{v} = 0 , \quad \nabla \cdot \boldsymbol{B} = 0 , \tag{1}$$

where we have set the density $\rho = 1$. In the presence of a constant background magnetic field $\boldsymbol{B} = B_0 \hat{z}$, the system of equations (1) admit exact solutions (in the ideal limit $\nu = \eta = 0$) representing Alfvénic disturbance propagating along $\hat{z}$ in either direction. These solutions are

$$\boldsymbol{B} = B_0 \hat{z} - \boldsymbol{f}^+(\boldsymbol{x}_\perp, z^+) , \quad \boldsymbol{v} = \boldsymbol{f}^+(\boldsymbol{x}_\perp, z^+) ,$$

and

$$\boldsymbol{B} = B_0 \hat{z} + \boldsymbol{f}^-(\boldsymbol{x}_\perp, z^-) , \quad \boldsymbol{v} = \boldsymbol{f}^-(\boldsymbol{x}_\perp, z^-) ,$$

where $\boldsymbol{x}_\perp \equiv (x, y)$, $z^\pm \equiv z \mp B_0 t$, and $\boldsymbol{f}^\pm(\boldsymbol{x})$ are arbitrary functions satisfying $\nabla \cdot \boldsymbol{f}^\pm = 0$. If $\boldsymbol{f}^\pm(\boldsymbol{x})$ are functions localized in z, then $\boldsymbol{f}^\pm(\boldsymbol{x}_\perp, z^\pm)$ represents Alfvén wave packets that propagate non-dispersively with the Alfvén speed $V_A = B_0$. These exact solutions propagate inward from $z = \mp\infty$ at $t \to -\infty$, retaining their form, until they collide. Because of the intrinsic nonlinearity of equation (1), the interaction of two colliding wave packets cannot be simply described by linear combinations of $\boldsymbol{f}^\pm(\boldsymbol{x}_\perp, z^\pm)$, for the linear combinations are not exact solutions of equation (1). In the weak-interaction limit (hereafter, referred to simply as the weak limit), we can write,

$$\boldsymbol{B} = B_0 \hat{z} - \boldsymbol{f}^+(\boldsymbol{x}_\perp, z^+) + \boldsymbol{f}^-(\boldsymbol{x}_\perp, z^-) + \boldsymbol{B}_1 , \qquad \boldsymbol{v} = \boldsymbol{f}^+(\boldsymbol{x}_\perp, z^+) + \boldsymbol{f}^-(\boldsymbol{x}_\perp, z^-) + \boldsymbol{v}_1 , \tag{2}$$

where $\boldsymbol{B}_1, \boldsymbol{v}_1$ are nonzero in general and are a measure the distortion of the wave packets when they collide. In the weak limit, as the magnitudes of $\boldsymbol{f}^\pm(\boldsymbol{x})$ (denoted by v) tend to zero, the magnitudes of $\boldsymbol{B}_1, \boldsymbol{v}_1$ (denoted by δv) tend to zero even faster. Following SG, we can estimate δv for the interactions between two Alfvén wave packets with typical size $(\kappa_\perp^{-1}, \kappa_z^{-1})$ by dimensional considerations. The collision time is of the order $(\kappa_z V_A)^{-1}$, and from equation (1),

$$\frac{dv}{dt} \sim \kappa_\perp v^2 , \qquad \frac{d^2 v}{dt^2} \sim \frac{d}{dt}(\kappa_\perp v^2) \sim \kappa_\perp^2 v^3 .$$

Then, for three-wave interactions,

$$\delta v \sim \frac{dv}{dt}(\kappa_z V_A)^{-1} \sim \left(\frac{\kappa_\perp v}{\kappa_z V_A}\right) v , \tag{3}$$

whereas, for four-wave interactions,

$$\delta v \sim \frac{d^2 v}{dt^2}(\kappa_z V_A)^{-2} \sim \left(\frac{\kappa_\perp v}{\kappa_z V_A}\right)^2 v . \tag{4}$$

We note that the main difference in scaling between three-wave and four-wave processes is that $\delta v \propto v^2$ for the former, whereas $\delta v \propto v^3$ for the latter. The k-dependence of the energy spectrum can be deduced if we consider MHD turbulence as a bath of mutually interacting Alvén wave packets. In the weak limit, it will take a large number of random collisions, $N \sim (v/\delta v)^2$, each with collision time $\tau \sim (\kappa_z V_A)^{-1}$, for a wave packet to lose memory of its initial state. If we assume that the cascade time of the energy is of the order $N\tau$, the rate of energy transfer per unit mass $\epsilon \sim v^2/N\tau$ is independent of k in the inertial range. Using $E(k)k \sim v^2$, we obtain

$$\epsilon \sim \frac{v^2}{N\tau} \sim kE\kappa_z V_A \left(\frac{\delta v}{v}\right)^2 \sim \text{constant} .$$

Using equation (3) and $\kappa_z \sim \kappa_\perp \sim k$, for three-wave interactions, we obtain the IK spectrum

$$E(k) \sim k^{-3/2} . \tag{5}$$

Note that we have assumed, for simplicity, that the spectrum is isotropic. While the issue of the isotropy of the spectrum is important and merits further investigation, it does not directly affect our discussion here which is primarily concerned with the question of whether three-wave or four-wave interactions dominate.

By equations (3) and (4), we see that the contributions from four-wave interactions are small compared with those from three-wave interactions, provided that v is small. However, if the resonant coupling coefficients for three-wave interactions are identically zero, as claimed by SG, then by equation (4), the energy spectrum, dominated by four-wave interactions, becomes

$$E(k) \sim k^{-4/3} . \tag{6}$$

Thus, the nature of the dominant resonant process underlying the interaction of Alfvén wave packets has a crucial effect on the exponent of k in the energy spectrum. In what follows, we study this basic process in detail, analytically as well as numerically.

3. THREE-WAVE INTERACTIONS IN IDEAL MHD

We consider the dynamical evolution of two localized Alfvén wave packets of the form equation (2) according to the ideal MHD equations, i.e., equations (1) with $\nu = \eta = 0$. In the weak limit, we can solve $\boldsymbol{B}_1, \boldsymbol{v}_1$ by perturbation theory. Substituting equation (2) into equation (1), and choosing the normalization $B_0 = 1$, we obtain

$$\frac{\partial \boldsymbol{v}_1}{\partial t} - \frac{\partial \boldsymbol{B}_1}{\partial z} = -2(\boldsymbol{f}^- \cdot \nabla \boldsymbol{f}^+ + \boldsymbol{f}^+ \cdot \nabla \boldsymbol{f}^-) - \nabla p , \qquad \frac{\partial \boldsymbol{B}_1}{\partial t} - \frac{\partial \boldsymbol{v}_1}{\partial z} = 2(\boldsymbol{f}^- \cdot \nabla \boldsymbol{f}^+ - \boldsymbol{f}^+ \cdot \nabla \boldsymbol{f}^-) . \tag{7}$$

We define $\boldsymbol{v}_1 \equiv (\boldsymbol{g}^- + \boldsymbol{g}^+)/2$, $\boldsymbol{B}_1 \equiv (\boldsymbol{g}^- - \boldsymbol{g}^+)/2$. Then, from equation (7), we obtain

$$\frac{\partial \boldsymbol{g}^+}{\partial t} + \frac{\partial \boldsymbol{g}^+}{\partial z} = -4\boldsymbol{f}^- \cdot \nabla \boldsymbol{f}^+ - \nabla p , \qquad \frac{\partial \boldsymbol{g}^-}{\partial t} - \frac{\partial \boldsymbol{g}^-}{\partial z} = -4\boldsymbol{f}^+ \cdot \nabla \boldsymbol{f}^- - \nabla p , \tag{8}$$

where p must satisfy the Poisson equation,

$$-\nabla^2 p = \nabla \cdot (4\boldsymbol{f}^- \cdot \nabla \boldsymbol{f}^+) = \nabla \cdot (4\boldsymbol{f}^+ \cdot \nabla \boldsymbol{f}^-) \tag{9}$$

in order that $\nabla \cdot \boldsymbol{g}^\pm = 0$.

We express all field variables in Fourier tansforms, e.g.,

$$\boldsymbol{g}^\pm(\boldsymbol{x}, t) = \int \tilde{\boldsymbol{g}}^\pm(\boldsymbol{k}, \omega) e^{i(\boldsymbol{k}\cdot\boldsymbol{x} - \omega t)}\, d\boldsymbol{k}\, d\omega . \tag{10}$$

Taking the Fourier transform of equation (8), we obtain

$$\pm i(k_z \mp \omega)\tilde{\boldsymbol{g}}^\pm(\boldsymbol{k}, \omega) = \tilde{\boldsymbol{F}}^\pm(\boldsymbol{k}, \omega) - \frac{\boldsymbol{k}\boldsymbol{k}}{k^2} \cdot \tilde{\boldsymbol{F}}^\pm(\boldsymbol{k}, \omega) \equiv \tilde{\boldsymbol{G}}^\pm(\boldsymbol{k}, \omega) , \tag{11}$$

where

$$\tilde{\boldsymbol{F}}^\pm(\boldsymbol{k}, \omega) = \frac{-4}{(2\pi)^4} \int \boldsymbol{f}^\mp(\boldsymbol{x}_\perp, z^\mp) \cdot \nabla \boldsymbol{f}^\pm(\boldsymbol{x}_\perp, z^\pm) e^{-i(\boldsymbol{k}\cdot\boldsymbol{x} - \omega t)}\, d\boldsymbol{x}\, dt . \tag{12}$$

We note that $z^\pm = z \mp t$. Defining $\kappa^\pm \equiv (k_z \pm \omega)/2$ and $\nabla^\pm \equiv \nabla_\perp + \hat{z}\partial/\partial z^\pm$, we have

$$\tilde{\boldsymbol{F}}^\pm = \frac{-2}{(2\pi)^2} \int e^{-i\boldsymbol{k}_\perp \cdot \boldsymbol{x}_\perp} d\boldsymbol{x}_\perp \left[\frac{1}{2\pi} \int \boldsymbol{f}^\mp(\boldsymbol{x}_\perp, z^\pm) e^{-i\kappa^\mp z^\mp}\, dz^\mp\right] \cdot \left[\frac{1}{2\pi} \int \nabla^\pm \boldsymbol{f}^\pm(\boldsymbol{x}_\perp, z^\pm) e^{-i\kappa^\pm z^\pm}\, dz^\pm\right]$$

$$\equiv \frac{-2}{(2\pi)^2} \int e^{-i\boldsymbol{k}_\perp \cdot \boldsymbol{x}_\perp} d\boldsymbol{x}_\perp\, \bar{\boldsymbol{f}}^\mp(\boldsymbol{x}_\perp, \kappa^\mp) \cdot [\nabla_\perp + i\kappa^\pm \hat{z}] \bar{\boldsymbol{f}}^\pm(\boldsymbol{x}_\perp, \kappa^\pm) . \tag{13}$$

Substituting equation (11) into equation (10), we solve for $\boldsymbol{g}^\pm(\boldsymbol{x}, t)$:

$$\boldsymbol{g}^\pm(\boldsymbol{x}, t) = \int \frac{\mp i}{\kappa^\mp} \tilde{\boldsymbol{G}}^\pm(\boldsymbol{k}, \omega) e^{i(\boldsymbol{k}_\perp \cdot \boldsymbol{x}_\perp + \kappa^+ z^+ + \kappa^- z^-)}\, d\boldsymbol{k}_\perp\, d\kappa^+ d\kappa^- . \tag{14}$$

As $t \to -\infty$, $\boldsymbol{f}^\pm(\boldsymbol{x}_\perp, z^\pm)$ represent two distinct incoming wave packets from $z \to \mp\infty$. There is no overlap between the wave packets in this limit, and hence causality requires that $\boldsymbol{g}^\pm(\boldsymbol{x}, t \to -\infty) \to 0$. The integration contours in the complex $\kappa^\pm$ plane have to be shifted $\pm 0^+$ from the real axis in order to satisfy this causality requirement. We are interested here in the asymptotic expressions of $\boldsymbol{g}^\pm(\boldsymbol{x}, t)$ as $t \to \infty$, which can be obtained from the contribution of the simple poles $\kappa^\pm = 0$ in equation (14). By Cauchy's residue theorem, we obtain

$$\boldsymbol{g}^\pm(\boldsymbol{x}, t \to \infty) \to 2\pi \int \tilde{\boldsymbol{G}}^\pm(\boldsymbol{k}, \pm k_z) e^{i(\boldsymbol{k}_\perp \cdot \boldsymbol{x}_\perp + k_z z^\pm)}\, d\boldsymbol{k}_\perp\, dk_z , \tag{15}$$

which represent two propagating Alfvén wave packets in the same directions as $f^{\pm}(x_\perp, z^{\pm})$. Now from equations (11) and (13), we have

$$\tilde{G}^{\pm}(\boldsymbol{k}, \pm k_z) = \tilde{F}^{\pm}(\boldsymbol{k}, \pm k_z) - \frac{\boldsymbol{k}\boldsymbol{k}}{k^2} \cdot \tilde{F}^{\pm}(\boldsymbol{k}, \pm k_z), \qquad \tilde{F}^{\pm}(\boldsymbol{k}, \pm k_z) = \frac{-2}{(2\pi)^2} \int e^{-i\boldsymbol{k}_\perp \cdot \boldsymbol{x}_\perp} d\boldsymbol{x}_\perp \bar{f}^{\mp}(\boldsymbol{x}_\perp, 0) \cdot [\nabla_\perp + ik_z \hat{z}] \bar{f}^{\pm}(\boldsymbol{x}_\perp, k_z). \tag{16}$$

The asymptotic solutions of $g^{\pm}$ are nonzero for arbitrary functional forms of $\bar{f}^{\pm}$. Note that $|g^{\pm}| \propto |f^{+}f^{-}|$ and hence, according to equation (3), is determined dominantly by three-wave interactions. We see that if $f^{\mp}(x_\perp, 0) = 0$, then $\tilde{G}^{\pm}(k, \pm k_z) = \tilde{F}(k_1 + k_z) = 0$, and consequently, $g^{\pm}(x, t \to \infty) \to 0$. In other words, the contribution from the resonant three-wave interactions vanishes if the $k_z = 0$ components of the Fourier transforms of the wave packets are zero. Since the $k_z = 0$ components of the Fourier transforms of the wave packets are generally nonzero in a turbulent bath, the three-wave interactions generally dominate.

To explicitly evaluate the contribution from four-wave interactions, we need to go to the second order in perturbation theory. In particular, we need to know the exact form of $g^{\pm}$ by equation (14), instead of the asymptotic expressions from equation (15). As shown below, the problem becomes analytically simpler using the RMHD equations.

4. ANALYTICAL CALCULATION OF THREE- AND FOUR-WAVE INTERACTIONS IN REDUCED MHD

In the presence of a large and uniform magnetic field $\boldsymbol{B} = \hat{z}$, the primitive MHD equations can be simplified if the plasma beta is low ($\beta \leq 1$). The primary effect of this large-scale field is to reduce the effective dimensionality of the incompressible MHD equations (Rosenbluth et al. 1976; Strauss 1976; Zank & Matthaeus 1993) because the spatial variation along the uniform magnetic field is much slower than variation transverse to the field. Even if the spectrum is initially isotropic, it becomes increasingly anisotropic with time (Montgomery 1982; Shebalin et al. 1983; Higdon 1984; SG). This legitimizes the use of RMHD ordering under which equation (1) can be simplified (in the ideal limit) to

$$\frac{\partial \Omega}{\partial t} - \frac{\partial J}{\partial z} = [A, J] - [\phi, \Omega], \qquad \frac{\partial A}{\partial t} - \frac{\partial \phi}{\partial z} = -[\phi, A], \tag{17}$$

where $\Omega = -\nabla_\perp^2 \phi$, $J = -\nabla_\perp^2 A$, $[\phi, A] \equiv \phi_y A_x - \phi_x A_y$, with $\boldsymbol{B} = \hat{z} + \nabla_\perp A \times \hat{z}$, $\boldsymbol{v} = \nabla_\perp \phi \times \hat{z}$. The RMHD equations (17) provide a useful description of nonlinear shear-Alfvén dynamics in the presence of a large-scale magnetic field. There is substantial observational rationale for the use of the system (17) as a model that captures the anisotropy of plasma turbulence in the ISM. Motivated by the observations of Armstrong, Cordes, & Rickett (1981) and the theoretical considerations of Montgomery (1982), Higdon (1984) has suggested that the anisotropy in ISM turbulence is a consequence of the presence of a large-scale, approximately uniform magnetic field in which density fluctuations are passively advected by the turbulent velocity field. Radio wave propagation observations give information on the anisotropy of the density irregularities. If the irregularities are isotropic, the blurring pattern or "seeing disk" of a source viewed through the turbulence is round and symmetric. On the other hand, if the irregularities are anisotropic, the blurring pattern is elongated. Radioastronomical observations of Spangler & Cordes (1988), and more recently Wilkinson, Narayan, & Spencer (1994), show that the turbulence in the ISM is anisotropic. From VLBI observations of Cyg X-3, Molnar et al. (1995) report that the scatering disk is anisotropic, with an axial ratio 1.31 ± 0.02.

The discussion above motivates us to explore nonlinear shear-Alfvén dynamics based on the RMHD equations. In analogy with equation (2), for weak interactions between two colliding shear-Alfvén wave packets, we seek perturbative solutions of the form

$$\phi = f^{-}(x_\perp, z^{-}) + f^{+}(x_\perp, z^{+}) + \phi_1 + \phi_2 + \cdots, \qquad A = f^{-}(x_\perp, z^{-}) - f^{+}(x_\perp, z^{+}) + A_1 + A_2 + \cdots. \tag{18}$$

The first-order solutions obey the equations,

$$\frac{\partial \Omega_1}{\partial t} - \frac{\partial J_1}{\partial z} = 2\{[f^{+}, \nabla_\perp^2 f^{-}] + [f^{-}, \nabla_\perp^2 f^{+}]\} \equiv F, \qquad \frac{\partial A_1}{\partial t} - \frac{\partial \phi_1}{\partial z} = 2[f^{-}, f^{+}] \equiv G. \tag{19}$$

As in § 2, the asymptotic expression of ϕ_1, A_1 are of the form,

$$\phi_1(x_\perp, t \to \infty) \to f_1^{-}(x_\perp, z^{-}) + f_1^{+}(x_\perp, z^{+}), \qquad A_1(x_\perp, t \to \infty) \to f_1^{-}(x_\perp, z^{-}) - f_1^{+}(x_\perp, z^{+}), \tag{20}$$

where

$$f_1^{\pm}(x_\perp, z) = \pi \int [\tilde{F}'(k, k_z) \mp \tilde{G}(k, \pm k_z)] e^{i\boldsymbol{k}\cdot\boldsymbol{x}} d\boldsymbol{k}, \tag{21}$$

has a form similar to equation (15) with $\tilde{F}'(k, \omega) \equiv \tilde{F}(k, \omega)/k_\perp^2$.

To simplify the derivation of the second-order solutions, we assume that the functions $f^{\pm}(x_\perp, z)$ are separable, i.e., $f^{\pm}(x_\perp, z) = f_\perp^{\pm}(x_\perp) f^{\pm}(z)$. The separability of $f^{\pm}(x_\perp, z)$ ensure that the functions F and G are also separable. Hence, we write

$$F(x, t) = F_\perp(x_\perp) f^{+}(z^{+}) f^{-}(z^{-}), \qquad G(x, t) = G_\perp(x_\perp) f^{+}(z) f^{-}(z),$$
$$\tilde{F}(k, \omega) = \tfrac{1}{2} \tilde{F}_\perp(k_\perp) \tilde{f}^{+}(\kappa^{+}) \tilde{f}^{-}(\kappa^{-}), \qquad \tilde{G}(k, \omega) = \tfrac{1}{2} \tilde{G}_\perp(k_\perp) \tilde{f}^{+}(\kappa^{+}) \tilde{f}^{-}(\kappa^{-}). \tag{22}$$

Then by equation (21),

$$f_1^{\pm}(\boldsymbol{x}_\perp, z) = \frac{\pi}{2}\int [\tilde{F}'_\perp \mp \tilde{G}_\perp] f^{\mp}(0)\tilde{f}^{\pm}(k_z)e^{i\boldsymbol{k}\cdot\boldsymbol{x}}\,d\boldsymbol{k}\,, \tag{23}$$

which is similar to equation (16). Once again, as in § 2, we conclude that the first-order perturbation is nonzero unless $\tilde{f}^{\pm}(0) = 0$. To find the exact functional forms of ϕ_1 and A_1, we define

$$\tilde{u}^{\pm} = \frac{1}{2}(\tilde{\phi}_1 \mp \tilde{A}_1) = \frac{\mp i(\tilde{F}' \mp \tilde{G})}{4\kappa^{\mp}}\,. \tag{24}$$

Then $u^{\pm}(\boldsymbol{x}, t)$ is given by

$$u^{\pm}(\boldsymbol{x}, t) = \int [\tilde{F}'_\perp \mp \tilde{G}_\perp]e^{i\boldsymbol{k}_\perp\cdot\boldsymbol{x}_\perp}\,d\boldsymbol{k}_\perp\left[\frac{\mp i}{4}\int \frac{d\kappa^{\mp}}{\kappa^{\mp}}\,\tilde{f}^{\mp}(\kappa^{\mp})e^{i\kappa^{\mp}z^{\mp}}\right] f^{\pm}(z^{\pm}) \equiv u_\perp^{\pm}(\boldsymbol{x}_\perp)\hat{f}^{\mp}(z^{\mp})f^{\pm}(z^{\pm})\,. \tag{25}$$

The complex integration in the complex $\kappa^{\pm}$ plane is carried out the same way as in § 2, and yields the asymptotic results,

$$\hat{f}^{\pm}(z) \to \begin{cases} 0 & \text{as} \quad z \to \pm\infty\,, \\ \pi\tilde{f}^{\pm}(0)/2 & \text{as} \quad z \to \mp\infty\,, \end{cases} \tag{26}$$

so that $u^{\pm}(\boldsymbol{x}, t \to \infty) \to f_1^{\pm}(\boldsymbol{x}_\perp, z^{\pm})$. We observe that the functions $f_1^{\pm}(\boldsymbol{x}_\perp, z^{\pm})$ have exactly the same z-dependence (via the functions $f^{\pm}[z^{\pm}]$) as the incoming wave packets $f^{\pm}(\boldsymbol{x}_\perp, z^{\pm})$. This implies that at first order there is no cascade of energy in the k_z direction, but only in the transverse direction, as exhibited by the form factors $F_\perp$ and $G_\perp$. Thus, the plasma becomes increasingly anisotropic and the RMHD ordering is strengthened further, making the RMHD approximation internally consistent (Montgomery 1982).

Now that we have explicit expressions (25) for the first-order solutions, we can calculate the second-order solutions. The relevant equations are

$$\begin{aligned} \frac{\partial\Omega_2}{\partial t} - \frac{\partial J_2}{\partial z} &= 2\{[f^+, \nabla_\perp^2 u^-] + [f^-, \nabla_\perp^2 u^+] + [u^+, \nabla_\perp^2 f^-] + [u^-, \nabla_\perp^2 f^+]\} \equiv H\,, \\ \frac{\partial A_2}{\partial d} - \frac{\partial\phi_2}{\partial z} &= 2\{[f^-, u^+] + [u^-, f^+]\} \equiv I\,. \end{aligned} \tag{27}$$

The function H has the following separable form:

$$\begin{aligned} H &= H_-(\boldsymbol{x}_\perp)f^-(z^-)\hat{f}^-(z^-)f^+(z^+) + H_+(\boldsymbol{x}_\perp)f^-(z^-)\hat{f}^+(z^+)f^+(z^+)\,, \\ \tilde{H}(\boldsymbol{k}, \omega) &= \frac{1}{2}\,\tilde{H}_-(\boldsymbol{k}_\perp)\left[\frac{1}{2\pi}\int f^-(z^-)\hat{f}^-(z^-)e^{-i\kappa^- z^-}\,dz^-\right]\tilde{f}^+(\kappa^+)\,, \\ &\quad + \frac{1}{2}\,\tilde{H}_+(\boldsymbol{k}_\perp)\left[\frac{1}{2\pi}\int f^+(z^+)\hat{f}^+(z^+)e^{-i\kappa^+ z^+}\,dz^+\right]\tilde{f}^-(\kappa^-)\,; \end{aligned} \tag{28}$$

similar expressions hold for I as well. As in equations (20) and (21) of the first-order calculation, asymptotic expressions for ϕ_2 and A_2 can be calculated from the relations

$$\phi_2(\boldsymbol{x}_\perp, t \to \infty) \to f_2^-(\boldsymbol{x}_\perp, z^-) + f_2^+(\boldsymbol{x}_\perp, z^+)\,, \qquad A_2(\boldsymbol{x}_\perp, t \to \infty) \to f_2^-(\boldsymbol{x}_\perp, z^-) - f_2^+(\boldsymbol{x}_\perp, z^+)\,,$$

$$f_2^{\pm}(\boldsymbol{x}_\perp, z) = \pi\int [\tilde{H}'(\boldsymbol{k}, \pm k_z) \mp \tilde{I}(\boldsymbol{k}, \pm k_z)]e^{i\boldsymbol{k}\cdot\boldsymbol{x}}\,d\boldsymbol{k}\,, \tag{29}$$

where $\tilde{H}' \equiv \tilde{H}/k_\perp^2$. Now, we consider

$$\tilde{H}(\boldsymbol{k}, \pm k_z) = \frac{\tilde{H}_\mp}{2}\left[\frac{1}{2\pi}\int f^{\mp}(z)\hat{f}^{\mp}(z)dz\right]\tilde{f}^{\pm}(k_z) + \frac{\tilde{H}_\pm}{2}\left[\frac{1}{2\pi}\int f^{\pm}(z)\hat{f}^{\pm}(z)e^{-ik_z z}\,dz\right]\tilde{f}^{\mp}(0)\,. \tag{30}$$

It can be shown that the first square bracket on the right-hand-side of equation (30) is equal to $\pi[\tilde{f}^{\mp}(0)]^2/4$. Substituting this result in equation (29), we obtain

$$f_2^{\pm}(\boldsymbol{x}_\perp, z) = [v^{\pm}(\boldsymbol{x}_\perp)\tilde{f}^{\mp}(0) + w^{\pm}(\boldsymbol{x}_\perp)\hat{f}^{\pm}(z)]\tilde{f}^{\pm}(0)f^{\pm}(z)\,, \tag{31}$$

where

$$v^{\pm}(\boldsymbol{x}_\perp) = \frac{\pi^2}{8}\int [\tilde{H}'_-(\boldsymbol{k}_\perp) \mp \tilde{I}_-(\boldsymbol{k}_\perp)]e^{i\boldsymbol{k}_\perp\cdot\boldsymbol{x}_\perp}d\boldsymbol{k}_\perp\,, \qquad w^{\pm}(\boldsymbol{x}_\perp) = \frac{\pi}{2}\int [\tilde{H}'_+(\boldsymbol{k}_\perp) \mp \tilde{I}_+(\boldsymbol{k}_\perp)]e^{i\boldsymbol{k}_\perp\cdot\boldsymbol{x}_\perp}\,d\boldsymbol{k}_\perp\,. \tag{32}$$

Since by equation (31), $|f_2| \propto |f|^3$, we see, upon comparison with equation (4), that the second-order perturbative solution has the scaling anticipated from four-wave interactions.

Equation (31) shows, somewhat surprisingly upon first glance, that for the case $\tilde{f}^{\pm}(0) = 0$, not only do the first-order (three-wave) asymptotic solutions $f_1^{\pm}$ vanish, but so do the second-order (four-wave) solutions $f_2^{\pm}$. However, the second-order solutions will not generally vanish if the wave packets are nonseparable.

We consider a simple example, with $f^{\pm}(z^{\pm}) = \cos(k^{\pm}z^{\pm})$, to illustrate clearly the physical content of the formal results obtained above. The cosine profiles do not represent wave packets in the usual sense because they are not spatially localized; rather, they may represent a Fourier component of a wave packet. By direct integration, we obtain

$$\tilde{F}(\boldsymbol{k}, \omega) = \tilde{F}(\boldsymbol{k}_\perp)[\delta(k_z - k^+ - k^-)\delta(\omega - k^+ + k^-) + \delta(k_z - k^+ + k^-)\delta(\omega - k^+ - k^-) + \delta(k_z + k^+ - k^-)\delta(\omega + k^+ + k^-) + \delta(k_z + k^+ + k^-)\delta(\omega + k^+ - k^-)]/8 . \quad (33)$$

In equation (33), the arguments of the δ-functions involving k_z and ω, respectively, represent the constraints of momentum and energy conservation. It is clear by inspection that the first-order solution can be attributed to three-wave interactions. We also see from equations (21) and (33) that $f_1^{\pm} = \tilde{F}(\boldsymbol{k}, \pm k_z) = 0$ except when $k^{\mp} = 0$ which means that if there a nonzero $k_z = 0$ component in the colliding wave packets, the first-order asymptotic solutions do not vanish. By equation (25), we can calculate the first-order solution exactly:

$$u^{\pm}(\boldsymbol{x}, t) = u_\perp^{\pm}(\boldsymbol{x}_\perp) \sin(k^{\mp}z^{\mp}) \cos(k^{\pm}z^{\pm}) . \quad (34)$$

We can see clearly from equation (33) that the three-wave interactions calculated here are resonant, and cannot be transformed away by a redefinition of variables (in the manner suggested by footnote 5 of SG). These interactions satisfy the energy and momentum conservation relations (required by the arguments of the delta functions) which are defining properties of resonant interactions (Zakharov, L'vov, & Falkovich 1992).

By the definition (28), we obtain

$$\begin{aligned}\tilde{H}(\boldsymbol{k}, \omega) = {} & \tilde{H}_-(\boldsymbol{k}_\perp)[\delta(k_z - k^+ - 2k^-)\delta(\omega - k^+ + 2k^-) - \delta(k_z - k^+ + 2k^-)\delta(\omega - k^+ - 2k^-) \\ & + \delta(k_z + k^+ - 2k^-)\delta(\omega + k^+ + 2k^-) - \delta(k_z + k^+ + 2k^-)\delta(\omega + k^+ - 2k^-)]/16i \\ & + \tilde{H}_+(\boldsymbol{k}_\perp)[\delta(k_z - 2k^+ - k^-)\delta(\omega - 2k^+ + k^-) + \delta(k_z - 2k^+ + k^-)\delta(\omega - 2k^+ - k^-) \\ & - \delta(k_z + 2k^+ - k^-)\delta(\omega + 2k^+ + k^-) - \delta(k_z + 2k^+ + k^-)\delta(\omega + 2k^+ - k^-)]/16 . \end{aligned} \quad (35)$$

By inspection of equation (35), it is clear that the second-order solution can be attributed to four-wave interactions. From equations (29) and (35), we also see that $f_2^{\pm} = \tilde{H}(\boldsymbol{k}, \pm k_z) = 0$ except when $k^{\mp} = 0$, which is similar to the property of the first-order solution. It can be inferred from this exercise that if the second-order asymptotic solutions are nonzero, then, in the weak limit, they are asymptotically much smaller than the first-order solutions (when the wave packets are separable). Note also that the second term in equation (31) or (35) indicates harmonic generation that may cause an energy cascade in k_z.

5. NUMERICAL RESULTS

Although the analytical results presented in § 4 are straightforward, it is useful to compare the predictions of the perturbative calculation with numerical results from the RMHD equations (17). We discuss the outcome of a simple numerical experiment in which two opposite propagating shear-Alfvén wave packets collide. We solve the RMHD equations numerically using periodic boundary conditions in the $\boldsymbol{x}_\perp$-plane and reflecting boundary condition in the z-direction. It is convenient to use a spectral decomposition in the $\boldsymbol{x}_\perp$-plane, and a finite-difference method along z. Since the wave packets are localized and do not reach the boundaries along z during simulations, the boundary conditions along z have a negligible effect on the results reported below.

The incoming wave packets used in the numerical calculations are of the form,

$$f_\perp^+(\boldsymbol{x}_\perp) = \sin(\kappa_\perp x) , \qquad f_\perp^-(\boldsymbol{x}_\perp) = \sin(\kappa_\perp y) , \qquad f^{\pm}(z) = c_\pm e^{-\kappa_z^2(z \pm z_0)^2} \cos[k_0(z \pm z_0) + a_\pm] , \quad (36)$$

where $c_\pm$, $a_\pm$, k_0, z_0, $\kappa_\perp$, and κ_z are real parameters. At $t = 0$, $f^{\pm}$ represent two localized wave packets centered at $z = \mp z_0$, propagating in the $\pm z$ direction. The first-order solution is given by equations (25) and (26). It can be shown easily that

$$u_\perp^{\pm}(\boldsymbol{x}_\perp) = \mp 2\kappa_\perp^2 \cos(\kappa_\perp x) \cos(\kappa_\perp y) , \qquad \tilde{f}^{\pm}(k) = \frac{c_\pm}{4\kappa_z\sqrt{\pi}} e^{\pm ikz_0}\{e^{[ia_\pm - (k \pm k_0)^2/4\kappa_z^2]} + e^{-ia_\pm - [(k \mp k_0)^2/4\kappa_z^2]}\} . \quad (37)$$

Hence, the $k_z = 0$ component of the wave packet is given by

$$\tilde{f}^{\pm}(0) = \frac{c_\pm \cos a_\pm}{2\kappa_z\sqrt{\pi}} e^{-(k_0^2/4\kappa_z^2)} , \quad (38)$$

which vanishes if $\cos a_\pm = 0$, and is very small when $k_0^2 \gg \kappa_z^2$. By equations (25), (26), and (36)–(38), we obtain

$$f_1^{\pm}(\boldsymbol{x}_\perp, z^{\pm}) = \frac{\mp c_+ c_- \sqrt{\pi}}{2\kappa_z} \kappa_\perp^2 \cos a_\mp \cos(\kappa_\perp x) \cos(\kappa_\perp y) e^{-(k_0^2/4\kappa_z^2) - \kappa_z^2(z^{\pm} \pm z_0)^2} \cos[k_0(z^{\pm} \pm z_0) + a_\pm] . \quad (39)$$

We note that the exact result equation (39) confirms the qualitative scaling (3) because $v \propto \kappa_\perp f$, $\delta v \propto \kappa_\perp f_1$.

To compensate for the effect of numerical dissipation, we find $f_1^{\pm}$ numerically by taking the difference of the run with initial conditions $f^{\pm}$ given by equation (36), and two other runs with the initial conditions f^+ or f^-. We have carried out runs with

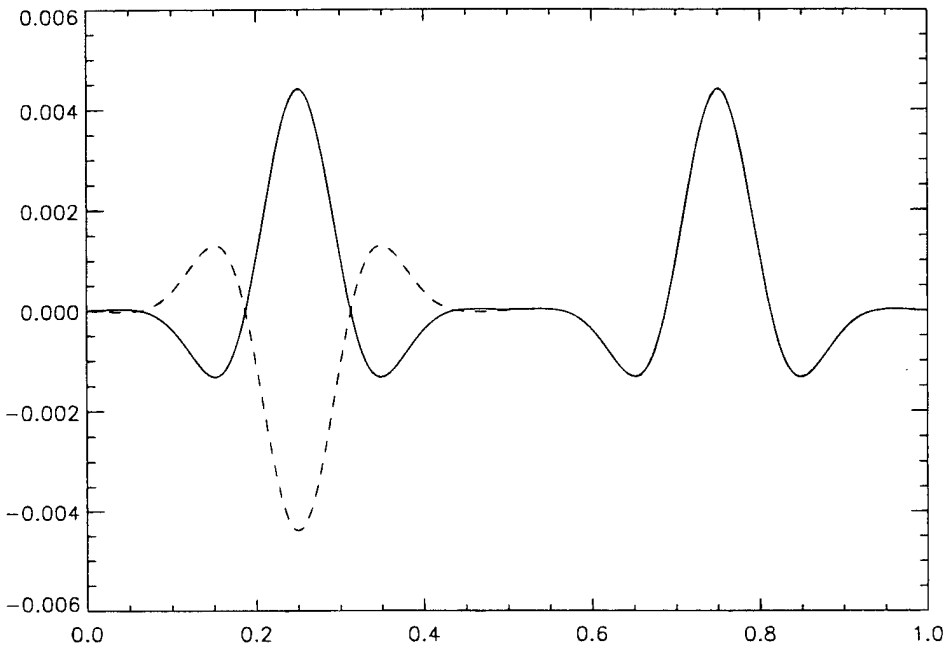

FIG. 1.—The initial condition at $t = 0$ for ϕ (*solid line*) and A (*dashed line*) at $x = y = 0.375$ along z, calculated by eq. (36) for parameters $\kappa_\perp = 2\pi$, $\kappa_z = 10$, $z_0 = 0.25$, $a_\pm = 0$, $k_0 = 8\pi$, $c_\pm = 0.0125$ and with spatial resolution $8 \times 8 \times 200$. The abscissae are $z + 0.5$.

the following parameters: $\kappa_\perp = 2\pi$ or 4π, $\kappa_z = 10$, 20 or 40, $z_0 = 0.25$, $a_\pm = 0$ or $\pi/2$, $k_0 = 0$, 8π, 16π or 24π, $c_\pm = 2''(0.00625)$ for $n = 1, 2, 3, \ldots$. Figure 1 shows the initial condition at $x = y = 0.375$ along z, for a run with $\kappa_\perp = 2\pi$, $\kappa_z = 10$, $z_0 = 0.25$, $a_\pm = 0$, $k_0 = 8\pi$, and $c_\pm = 0.0125$. The results for the spatio-temporal evolution agree with equation (39) for most cases with errors as small as 0.4% and rarely larger than 10%. We have verified the functional forms in x, y, and z, as well as the parametric dependence on $\kappa_\perp$, κ_z, $c_\pm$ and k_0. For the initial condition shown in Figure 1, Figure 2 shows the comparison between analytical and numerical results for the first order fields at $t = 0.5$, along z for $x = y = 0.375$. The difference between the analytical and numerical results increases when $f_1^\pm$ is small or $c_\pm$ is large. For example, when $\cos a_\pm = 0$ (or $k_0^2 \gg \kappa_z^2$), equation (39) predicts that $f_1^\pm$ should be zero (or extremely small), and hence, there should be no (or extremely small) first-order asymptotic fields. However, in the numerical experiment, the presence of numerical error fields ($\sim 10^{-8}$) leaves first-order residues that are somewhat larger than the analytical predictions.

Although the first-order solution is, in general, much larger than the second-order solution, we see from equation (39) that when $\cos(\kappa_\perp x) \cos(\kappa_\perp y) = 0$, the first-order solution vanishes. Along these lines in $x - y$ space, the second-order fields can be

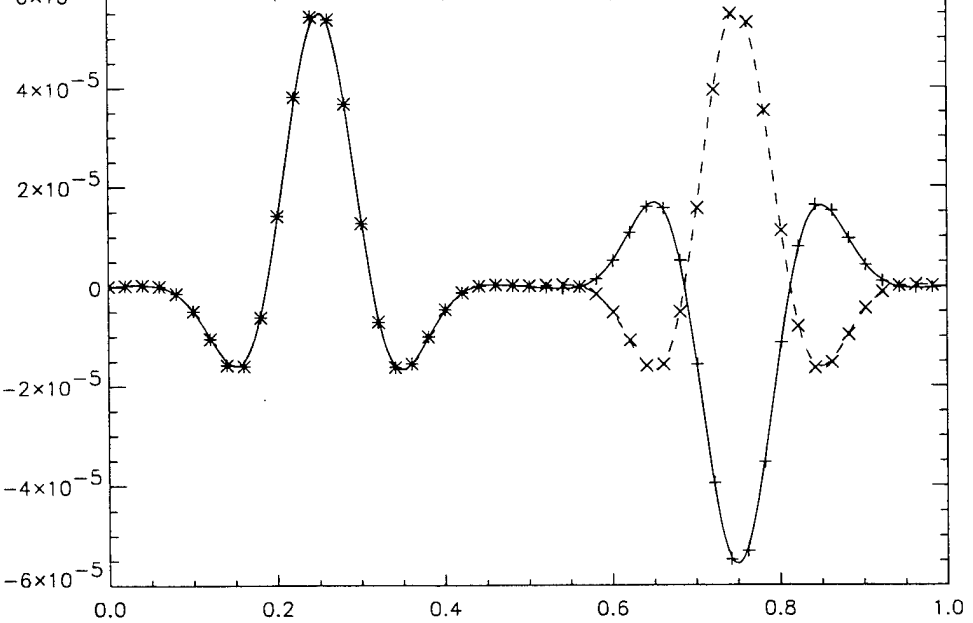

FIG. 2.—First-order fields ϕ_1 (*solid line*) and A_1 (*dashed line*) from the simulation at $t = 0.5$ for the initial condition of Fig. 1, along z for $x = y = 0.375$. Comparison with $\phi_1(+)$ and $A_1(\times)$, given analytically by eqs. (20) and (39).

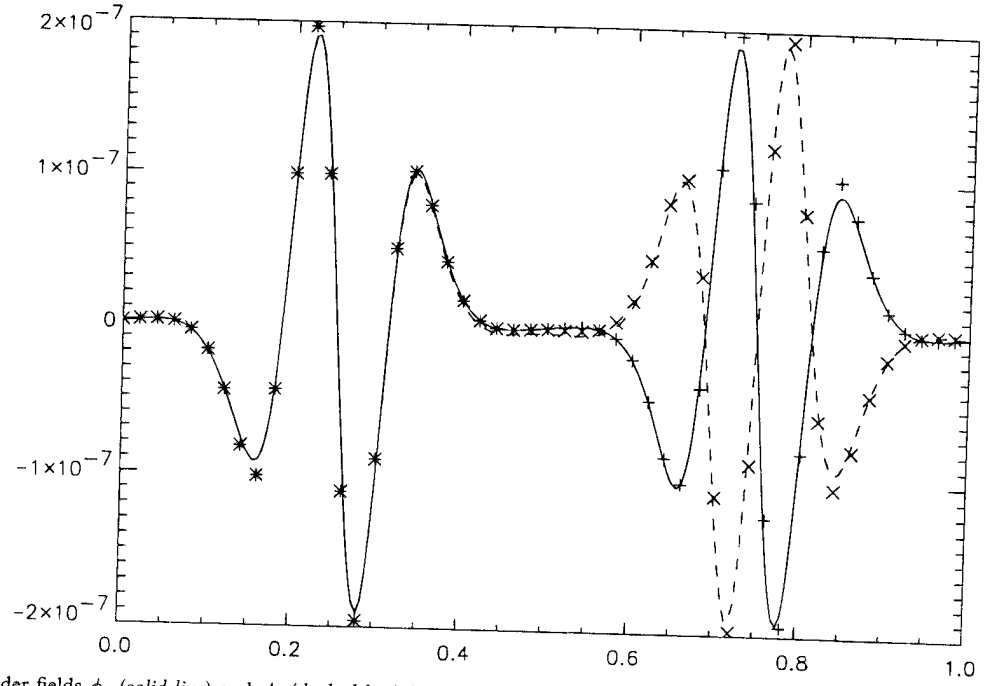

FIG. 3.—Second-order fields ϕ_2 (*solid line*) and A_2 (*dashed line*) found by simulation at $t = 0.5$ for the initial condition of Fig. 1, along z for $x = 0.25$, $y = 0.75$. Comparison with $\phi_2(+)$ and $A_2(\times)$, given analytically by eqs. (29), (31), (36), (40), and (41).

found numerically and their values compared with the analytical predictions (31) and (32). For the initial conditions (36), we obtain

$$v^{\pm}(\boldsymbol{x}_\perp) = -\frac{\pi^2}{2}\kappa_\perp^4 \sin(\kappa_\perp \xi^{\pm})\left[\frac{3}{5}\cos(2\kappa_\perp \xi^{\mp}) + 1\right], \qquad w^{\pm}(\boldsymbol{x}_\perp) = \frac{4\pi}{5}\kappa_\perp^4 \sin(\kappa_\perp \xi^{\mp})\cos(2\kappa_\perp \xi^{\pm}), \tag{40}$$

where $\xi^+ = x$, $\xi^- = y$. In general, the expressions of $f_2^{\pm}$ in equation (31) contain the functions $\hat{f}^{\pm}(z)$ which are defined in equation (25) by nontrivial contour integrals. However, with $f^{\pm}(z)$ given by equation (36), $\hat{f}^{\pm}(z)$ can be obtained by direct integration. We obtain

$$\hat{f}^{\pm}(z) = \mathrm{Re}\,\frac{c_\pm \sqrt{\pi}}{8\kappa_z}\left\{1 \mp \mathrm{erf}\left[\kappa_z\left(z \pm z_0 \pm \frac{ik_0}{2\kappa_z^2}\right)\right]\right\} e^{-(k_0^2/4\kappa_z^2) + ia_\pm}, \tag{41}$$

where erf denotes the error function and Re denotes taking the real part. In particular, $\hat{f}^{\pm}(\mp z_0) = c_\pm \pi^{1/2} \exp(-k_0^2/4\kappa_z^2)/8\kappa_z$. Hence, we can easily compare the analytical results equation (31) with the numerical results, using equations (36), (38), (40) and (41) at the points $z^{\pm} = \mp z_0$. We find reasonably good agreement, with error smaller than 10% in most cases, even though the absolute values of the second-order solutions are very small (down to 10^{-8}). The functional forms of the solutions are qualitatively consistent with the shape of the functions $\hat{f}^{\pm}(z)$. The numerical scaling with $\kappa_\perp$, κ_z, $c_\pm$ also agree with analytical prediction and are consistent with equation (4). Figure 3 shows the numerical and analytical second-order fields at $t = 0.5$ for the initial condition of Figure 1, along z for $x = 0.25$, $y = 0.75$. The second-order fields, which are expected to vanish (or be very small) for $\tilde{f}^{\pm}(0) = 0$ when $\cos a_\pm = 0$ (or $k_0^2 \gg \kappa_z^2$) are also dominated by error fields.

In summary, the analytical predictions from first and second-order perturbation theory are in good accord with the RMHD simulations.

6. CONCLUSIONS

We have demonstrated analytically as well as numerically that the weak interactions between Alfvén wave packets are, in general, dominated by three-wave interactions except in the special cases in which the $k_z = 0$ components of the Fourier transforms of the packets are zero. This assertion holds for both the MHD and RMHD equations. Since the three-wave interactions are dominant in the weak limit, the exponent of the wavenumber in the energy spectrum is given by the IK theory if the heuristic arguments in § 2 are valid. The theory of SG associates no energy with the $k_z = 0$ components of the Fourier transforms of the packets, as discussed clearly in the critique of Montgomery & Matthaeus (1995). At a technical level, SG's elegant variational analysis omits third-order terms (proportional to the cube of the virtual displacement) in the Lagrangian, and this omission leads to the conclusion that the three-wave interactions are empty. The technical problems involved in SG's analysis are discussed in more detail in the Appendix.

Although motivated in part by SG's critique of the IK theory, this paper is more than a critique of SG. What we have attempted to do is to develop a perturbative method for the explicit determination of three- and four-wave interactions in weak MHD turbulence in the presence of a strong magnetic field. A better understanding of these interactions may be helpful in developing a more complete analytical theory of weak and strong MHD turbulence and in avoiding possible pitfalls.

While we have provided fundamental arguments in support of the general dominance of the three-wave interactions underlying the IK theory, it would be premature to suggest that the IK theory is a valid theory of fluid turbulence in the ISM. Besides the questionable assumption of isotropy, effects of compressibility will lead to nonlinear steepening of Alfvén waves in a manner excluded by the incompressible theory. Collisionless effects, not included in the Ohm's law of resistive MHD, are likely to have a profound effect on the nature of the turbulent dynamics. Nonetheless, incompressible MHD or even RMHD is a useful starting point, because it allows us to focus on aspects of shear-Alvén dynamics that have fundamental implications for problems of astrophysical plasma turbulence.

We are grateful to S. Spangler for helpful discussions on the interpretation of radioastronomical observations. We thank P. Goldreich for a useful and constructive exchange of ideas; the contents of the Appendix were stimulated by this exchange. This research is supported by the NSF grant No. ATM 93-10157 and the AFOSR grant No. F49620-93-1-0071. Supercomputing resources were provided by the San Diego Supercomputing Center.

APPENDIX A

REMARKS ON SG'S LAGRANGIAN ANALYSIS

We have shown that the three-wave interactions are generally nonzero unless the $k_z = 0$ component vanishes. Therefore, SG's conclusion that three-wave interactions always vanish can be traced to the incorrect assumption that there is no $k_z = 0$ component, or in their own words, such "waves possess no power to contribute to resonant interactions." The physical flaw underlying SG's reasoning, as pointed out by Montgomery & Matthaeus (1995), is that such a component cannot be treated as a wave, as distinct from a spatial Fourier mode. Here we point out the technical problems SG's Lagrangian treatment.

SG begin from the Lagrangian

$$L = \int d^3x \left(\frac{\rho v^2}{2} - \frac{\boldsymbol{B}^2}{8\pi} \right), \tag{A1}$$

for the ideal incompressible MHD model with constant density ρ. Following Newcomb (1962), they treat the position vector $\boldsymbol{x}$ as the Lagrangian variable of a fluid element such that

$$\boldsymbol{x} = \boldsymbol{x}_0 + \xi(\boldsymbol{x}_0, t) , \tag{A2}$$

where ξ is the displacement vector following the fluid element with initial position $\boldsymbol{x}_0 + \xi(\boldsymbol{x}_0, 0)$. Then the velocity $\boldsymbol{v}$ and magnetic field $\boldsymbol{B}$ of the fluid element can be written as

$$\boldsymbol{v} = \frac{\partial \xi(\boldsymbol{x}_0, t)}{\partial t} , \qquad \boldsymbol{B} = \boldsymbol{B}_0 + \boldsymbol{B}_0 \cdot \frac{\partial \xi(\boldsymbol{x}_0, t)}{\partial \boldsymbol{x}_0} , \tag{A3}$$

where $\boldsymbol{B}_0$ can be chosen to be the constant magnetic field $B_0 \hat{z}$. For small ξ, using equation (13), we can write the $k_z = 0$ component of $\boldsymbol{B}$ to the lowest order in ξ, as

$$\bar{\boldsymbol{B}}(\boldsymbol{x}_\perp, k_z = 0) \approx B_0 \int \frac{\partial \xi}{\partial z} \, dz = B_0 [\xi(\boldsymbol{x}_\perp, \infty) - \xi(\boldsymbol{x}_\perp, -\infty)] , \tag{A4}$$

which is in general nonzero, unless we make one of two possible asumptions: either $\xi(\boldsymbol{x}_\perp, \infty) = \xi(\boldsymbol{x}_\perp, -\infty)$ or $\xi(\boldsymbol{x}_\perp, z \to \pm\infty) \to 0$. The first assumption is too special to be of general interest, and so we consider the second possibility. Even is we require the Alfvén wave packets to be so localized that $\boldsymbol{v}, \boldsymbol{B} \to 0$ as $z \to \pm\infty$, it is not necessary that ξ should vanish as $z \to \pm\infty$ because the fields $\boldsymbol{v}, \boldsymbol{B}$ are expressed as derivatives of ξ by equation (A3). For example, if we use the wave packets (36), which is a legitimate choice with $\boldsymbol{v}, \boldsymbol{B} \to 0$ as $z \to \pm\infty$, and calculate ξ, we will not generally obtain ξ as $z \to \pm\infty$. Hence, the second possibility is overrestrictive. We conclude that the right-hand side of equation (A4) is nonzero in general. The appropriate boundary condition for the displacement vector is $\xi(\boldsymbol{x}_\perp, z \to \pm\infty) \to \xi^{\pm}(\boldsymbol{x}_\perp)$ with $\xi^{+}(\boldsymbol{x}_\perp) \neq \xi^{-}(\boldsymbol{x}_\perp) \neq 0$ in general. Because of this, the definition of the Fourier transform of ξ becomes problematic, and may not even exist if, in particular, $\bar{\xi}(\boldsymbol{x}_\perp, k_z \to 0) \to \infty$.

Without addressing the question of existence of the Fourier transform of ξ, SG write the following Lagrangian L by integrating over k-space:

$$L(\xi) = \frac{\rho}{2} \int \frac{d^3k}{8\pi^3} \left[\dot{\bar{\xi}}(\boldsymbol{k}) \cdot \dot{\bar{\xi}}(-\boldsymbol{k}) - \omega^2(\boldsymbol{k}) \bar{\xi}(\boldsymbol{k}) \cdot \bar{\xi}(-\boldsymbol{k}) \right], \tag{A5}$$

where $\omega(\boldsymbol{k}) = V_A |k_z|$ and $V_A = B_0/(4\pi\rho)^{1/2}$ being the Alfvén speed. Note that the Fourier transforms are taken in $\boldsymbol{x}_0$-space ("unperturbed space") but not the $\boldsymbol{x}$-space ("perturbed space"), with $\boldsymbol{x}$ related to $\boldsymbol{x}_0$ by equation (A2). The first variation with respect to ξ must satisfy the Jacobian constraint

$$J \equiv \left\| \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{x}_0} \right\| = \left\| \delta_{ij} + \frac{\partial \xi_i}{\partial x_{0j}} \right\| = 1 , \tag{A6}$$

for incompressible plasmas. This constraint can be incorporated by using either a Lagrange multiplier or by solving equation (A6) directly and substituting the results in $L(\xi)$. SG implement the latter procedure perturbatively to lowest order in ξ and obtain

$$L = L(\eta) + \mathcal{O}(\eta^4) . \tag{A7}$$

where $L(\eta)$ is defined by substituting η for ξ in equation (A5) and $\tilde{\xi}_i(\boldsymbol{k}) = \tilde{\eta}_i(\boldsymbol{k}) + k_i\tilde{\phi}(\boldsymbol{k})/k$ with $k_i\tilde{\eta}_i(\boldsymbol{k}) = 0$. Since there is no η^3 term in equation (A7), SG conclude that three-wave interactions are empty. If this conclusion is indeed correct, then the total Lagrangian $L(\xi)$ given by equation (A5) with $\nabla_0 \cdot \xi = 0$ should, to this order, describe a system of noninteracting waves. But as we show below, this is not so. (In the weak limit, omitting four-wave and other higher order terms, we can omit the Jacobian constraint.) The Euler-Lagrange equation is

$$\ddot{\tilde{\xi}}(\boldsymbol{k}) + \omega^2(\boldsymbol{k})\tilde{\xi}(\boldsymbol{k}) = 0 , \tag{A8}$$

which has the general solution

$$\tilde{\xi}(\boldsymbol{k}, t) = \tilde{\xi}_1(\boldsymbol{k})e^{i\omega(\boldsymbol{k})t} + \tilde{\xi}_2(\boldsymbol{k})e^{-i\omega(\boldsymbol{k})t} . \tag{A9}$$

Transforming to real x_0-space, we obtain

$$\xi(\boldsymbol{x}_0, t) = \xi_1(\boldsymbol{x}_{\perp 0}, z_0 + V_A t) + \xi_2(\boldsymbol{x}_{\perp 0}, z_0 - V_A t) . \tag{A10}$$

Using equation (A3), the $\boldsymbol{v}$, $\boldsymbol{B}$ fields of the wave packets can be expressed as

$$\boldsymbol{v}(\boldsymbol{x}_0, t) = \boldsymbol{v}_1(\boldsymbol{x}_{\perp 0}, z_0 + V_A t) - \boldsymbol{v}_2(\boldsymbol{x}_{\perp 0}, z_0 - V_A t) \qquad \boldsymbol{B}(\boldsymbol{x}_0, t) = \boldsymbol{B}_1(\boldsymbol{x}_{\perp 0}, z_0 + V_A t) + \boldsymbol{B}_2(\boldsymbol{x}_{\perp 0}, z_0 - V_A t) , \tag{A11}$$

with $\boldsymbol{B}_i = B_0\boldsymbol{v}_i/V_A$. Upon first glance, it may appear that the two localized Alfvén wave packets propagating in opposite directions have no interaction at all. However, here $\boldsymbol{B}_i$ and $\boldsymbol{v}_i$ are expressed as functions of $\boldsymbol{x}_0$, as distinct from $\boldsymbol{f}^{\pm}$ of equation (2) which are functions of $\boldsymbol{x}$. Expressing the fields as functions of $\boldsymbol{x}$ by means of the relation (A2), we find that the perturbation terms $\boldsymbol{B}_1, \boldsymbol{v}_1$ in equation (2) are not zero. Indeed, to lowest order in ξ as $t \to \infty$, we obtain

$$\Delta\boldsymbol{B}_1(\boldsymbol{x}_\perp, z + V_A t) \approx -\bar{\boldsymbol{B}}_2(\boldsymbol{x}_\perp, k_z = 0) \cdot \nabla\boldsymbol{B}_1(\boldsymbol{x}_{\perp 0}, z_0 + V_A t)/B_0 , \tag{A12}$$

using equation (A4). Note that equation (A12) is identical to the result obtained earlier using equations (15) and (16) to the same order of ξ, based on three-wave interactions. Contrary to SG's claim, we have thus shown that in the "perturbed space," there are nonvanishing three-wave interactions in the Lagrangian (A7) even though there are no $\mathcal{O}(\eta^3)$ terms. This becomes even more transparent if we follow SG and take Fourier transforms but do so in the "perturbed space" at the outset, in which case the total Lagrangian will contain manifestly nonzero third-order terms. For example, using the transformation

$$\left.\frac{\partial\xi}{\partial t}\right|_{x_0} = \left.\frac{\partial\xi}{\partial t}\right|_{x} + \left.\frac{\partial\xi}{\partial \boldsymbol{x}}\right|_{t} \cdot \left.\frac{\partial\xi}{\partial t}\right|_{x_0} , \tag{A13}$$

iteratively, the first term in the integrand of L, to third order ξ, is given by

$$\int \dot{\xi}^2\, d\boldsymbol{x} = \frac{1}{(2\pi)^3}\int \dot{\tilde{\xi}}(\boldsymbol{k}) \cdot \dot{\tilde{\xi}}(-\boldsymbol{k})d\boldsymbol{k} + \frac{2i}{(2\pi)^3}\int \dot{\tilde{\xi}}(\boldsymbol{k}) \cdot \tilde{\xi}(\boldsymbol{k}')\boldsymbol{k}' \cdot \dot{\tilde{\xi}}(\boldsymbol{k}'')\delta(\boldsymbol{k} + \boldsymbol{k}' + \boldsymbol{k}'')d\boldsymbol{k}\, d\boldsymbol{k}'\, d\boldsymbol{k}'' . \tag{A14}$$

While this suggests qualitatively why our conclusions differ from SG's, the derivation of the correct third-order Lagrangian with general boundary conditions is beyond the scope of this paper.

REFERENCES

Armstrong, J. W., Cordes, J. M., & Rickett, B. J. 1981, Nature, 291, 561
Biskamp, D. 1993, Nonlinear Magnetohydrodynamics (Cambridge: Cambridge Univ. Press)
Goldreich, P., & Sridhar, S. 1995, ApJ, 438, 763
Higdon, J. C. 1984, ApJ, 285, 109
Iroshnikov, P. S. 1963, AZh, 49, 742
Kraichnan, R. H. 1965, Phys. Fluids, 8, 1385
Molnar, L. A., Mutel, R. L., Reid, M. J., & Johnston, K. J. 1995, ApJ, 438, 708
Montgomery, D. 1982, Phys. Scripta, T2, 83
Montgomery, D., & Matthaeus, W. H. 1995, ApJ, 447, 706
Narayan, R. 1992, Phil. Trans. R. Soc. Lond. A, 341, 151
Newcomb, S. A. 1962, Nucl. Fusion Suppl., Part 2, 451
Rickett, B. J. 1990, ARA&A, 28, 561
Rosenbluth, M. N., Monticello, D. A., Strauss, H. R., & White, R. B. 1976, Phys. Fluids, 19, 1987
Shebalin, J. V., Matthaeus, W. H., & Montgomery, D. 1983, J. Plasma Phys., 29, 525
Spangler, S. R., & Cordes, J. M. 1988, ApJ, 332, 346
Sridhar, S., & Goldreich, P. 1994, ApJ, 432, 612 (SG)
Strauss, H. R. 1976, Phys. Fluids, 19, 134
Wilkinson, P. N., Narayan, R., & Spencer, R. E. 1994, MNRAS, 269, 67
Zakharov, V. E., L'vov, V. S., Falkovich, G. 1992, Kolmogorov Spectra of Turbulence I (Berlin: Springer)
Zank, G. P., & Matthaeus, W. H. 1993, Phys. Fluids, A5, 257

Scaling of anisotropic spectra due to the weak interaction of shear-Alfvén wave packets

C. S. Ng and A. Bhattacharjee
Department of Physics and Astronomy, The University of Iowa, Iowa City, Iowa 52242

(Received 22 July 1996; accepted 18 November 1996)

Weak magnetohydrodynamic turbulence in the presence of a uniform magnetic field is dominated by three-wave interactions that mediate the collisions of shear-Alfvén wave packets propagating in opposite directions parallel to the magnetic field. The scaling of three-wave couplings is calculated by asymptotic analysis and a direct numerical evaluation of the nonlinear interaction based on the reduced magnetohydrodynamic equations. A new relation is derived between the spectral index of three-wave coupling and the spectral indices of two random-amplitude wave packets. This relation has significant implications for the anisotropic energy spectrum. © *1997 American Institute of Physics.* [S1070-664X(97)00803-3]

I. INTRODUCTION

The structure and scaling of magnetohydrodynamic (MHD) turbulence are issues of considerable interest for laboratory as well as astrophysical plasmas. Incompressible magnetohydrodynamics, despite its limitations as a model, is usually the starting point for the investigation of these issues. In the presence of a directed magnetic field, MHD turbulence tends to exhibit a pronounced anisotropy. The main goal of this paper is to examine, in the limit of weak turbulence, the nature and scaling of the anisotropic fluctuations in a plasma permeated by a spatially uniform magnetic field $\mathbf{B}=B_0\hat{\mathbf{z}}$. We demonstrate that the scaling of the anisotropic fluctuations even in this highly simplified configuration exhibits features that raise serious questions regarding the validity of some common assumptions made in scaling or closure studies of MHD turbulence.

In view of the striking success of Kolmogorov's original theory of hydrodynamic (HD) turbulence,[1] it is plausible that early work in MHD turbulence relied on some of the same assumptions as the Kolmogorov theory after incorporating new dynamical features introduced by the presence of magnetic fields. Two such well-known assumptions are the *isotropy* of the turbulence and the dominance of *local interactions* in k-space. In the Iroshnikov–Kraichnan (hereafter, IK) theory of MHD turbulence,[2,3] small-scale fluctuations are envisioned to behave as Alfvén wave packets propagating along magnetic lines of force, and the collision between two oppositely propagating wave packets provides a mechanism for the cascade of energy to short wavelengths. This picture of colliding wave packets mediating the cascade to short wavelengths not only provides physical insight, but is also supported by evidence in numerical experiments[4,5] as well as *in situ* observations of the solar wind[6–9] that Alfvénic fluctuations play an important role in the energy transfer between the fluid and magnetic field fluctuations. While the IK theory elucidates the role of colliding Alfvén wave packets in the generation of the cascade, it retains the assumptions of isotropy and local interactions underlying the Kolmogorov theory.

The violation of the assumption of isotropy in the presence of a mean magnetic field has been recognized by several investigators.[10–17] However, the assumption of local interactions has not received the same amount of scrutiny. One of the objectives of this paper is to revisit this assumption, and provide analytical and numerical evidence that raise questions about its validity. We also show that the violation of this assumption can have a significant effect on the scaling of the anisotropic energy spectrum.

II. DIMENSIONAL AND HEURISTIC ANALYSIS

In order to elucidate the important role that the assumptions of isotropy and local interactions have played in theories of MHD turbulence, we begin with a brief review of scaling arguments in HD as well as MHD. Kolmogorov[1] demonstrated that the turbulent energy spectrum in the inertial range of an incompressible fluid can be found by dimensional analysis. His argument can be summarized as follows. Let us assume that the turbulence is *isotropic* in wave number (k) space. Then the total energy can be written $\int E(k)dk$ where $E(k)$ is the turbulent energy spectrum. Assume that there exists an inertial range such that the energy transfer rate $\varepsilon(k)$ is constant (independent of k) and, furthermore, that the energy transfer is *local* in k-space. Since the dimension of ε is L^2T^{-3}, k is L^{-1} and $E(k)$ is L^3T^{-2}, where L is the dimension of length and T is of time, dimensional homogeneity of the relation $\varepsilon \sim k^{\alpha}E_k^{\beta}$ yields the inertial-range energy spectrum $E(k) \propto \varepsilon^{2/3}k^{-5/3}$.

Iroshnikov[2] and Kraichnan[3] extended Kolmogorov's analysis to incompressible magnetohydrodynamic (MHD) turbulence. As discussed by Kraichnan,[3] there is a crucial distinction between hydrodynamic (HD) and MHD flows: whereas a uniform velocity in a HD flow can be transformed away by a Galilean transformation, a uniform magnetic field in a MHD flow cannot. The imposition of a background magnetic field causes wave disturbances to travel along the uniform magnetic field in both directions with Alfvén speed V_A (which is a measure of the magnetic field strength). A wave packet can interact with another wave packet only if the two are moving in opposite directions and collide. Since the interaction time $\tau \sim (kV_A)^{-1}$ is usually much shorter than the eddy turn over time $k^{-1}v(k)^{-1}$, where $v(k)$ is the typical amplitude of the wave (in dimension of velocity), the

 1070-664X/97/4(3)/605/6/$10.00 © 1997 American Institute of Physics

energy cascade is more inhibited in MHD than it is in HD. By treating the $k=0$ component of the magnetic field at any spatial location as the locally uniform field, and assuming that the MHD turbulence is isotropic and local in k-space, Kolmogorov's dimensional arguments can be repeated, now with ε depending on k, $E(k)$, and V_A (with dimension LT^{-1}). Writing $\varepsilon \sim k^{\alpha} E_k^{\beta} V_A^{\gamma}$, we can deduce the spectral index ν of the inertial-range energy spectrum:

$$\nu = \frac{\alpha}{\beta} = \frac{5-\gamma}{3-\gamma}, \tag{1}$$

with the unknown γ to be determined. [Note that the Kolmogorov spectrum $\nu = \frac{5}{3}$ is obtained if we set $\gamma = 0$ in Eq. (1).] To find ν, we must determine the value of γ which gives the power law dependence of ε on V_A. This can be done in the limit of weak turbulence when the interaction between two wave packets can be treated by perturbation theory. The lowest-order interaction involves three-wave interactions during which two wave packets collide for a typical time scale $\tau \sim (kV_A)^{-1}$ and produce a third wave with typical velocity magnitude δv. By the relation $\varepsilon \propto (\delta v)^2/\tau \propto \tau$, we obtain $\gamma = -1$ and thus $\nu = \frac{3}{2}$, which yields the Iroshnikov–Kraichnan (IK) spectrum $E(k) \propto \varepsilon^{1/2} k^{-3/2}$.

The scaling results obtained above by dimensional analysis for isotropic MHD turbulence can also be obtained by an alternate physical argument.[13,14] Let each of the two colliding wave packets have amplitude of the order v and spatial scale k^{-1}. Assume that the energy transfer is local in k space, which means that a wave will interact dominantly with another wave with the same length scale but moving in the opposite direction. From the MHD equations, we estimate that $\dot{v} \sim kv^2$, where an overdot denotes time derivative. Then, if three-wave interactions dominate, we can write $\delta v \sim \dot{v}\tau \sim v^2/V_A$. In the weak limit, it will take a large number of random collisions, $N \sim (v/\delta v)^2 \gg 1$, to change a wave packet amplitude significantly. Assume that the energy content of the wave in this length scale cascade to the smaller scale after N collisions. Hence, we obtain $\varepsilon \sim v^2/N\tau \sim k^2 E(k) V_A / N \sim k^3 E^2(k) V_A^{-1}$, which implies that $E(k) \propto \varepsilon^{1/2} k^{-3/2}$, also obtained by dimensional analysis.

The physical argument discussed in the preceding paragraph shows how a heuristic scaling can be obtained when dimensional analysis fails. Dimensional analysis is useful in determining the spectral index if the turbulence is isotropic. However, in the presence of a background magnetic field, the MHD turbulence is anisotropic.[10–17] Dimensional analysis cannot then provide a definite result since it cannot discriminate between the two length scales perpendicular ($k_\perp^{-1}$) and parallel ($k_\parallel^{-1}$) to the uniform magnetic field. Let us assume that the energy cascade occurs entirely in the direction perpendicular to the uniform field so that the total energy can be written $\int E(k_\perp) dk_\perp dk_\parallel$. If we now repeat the scaling argument given in the last paragraph with $\tau \sim (k_\parallel V_A)^{-1}$, we obtain

$$\delta v \sim k_\perp v^2 / k_\parallel V_A \sim k_\perp^2 E(k_\perp)/V_A. \tag{2}$$

Thus $\varepsilon \sim k_\perp^4 k_\parallel E^2(k_\perp)/V_A$, which implies that the anisotropic spectrum is $E(k_\perp) \propto \varepsilon^{1/2} k_\perp^{-2}$ for weak MHD turbulence dominated by three-wave interactions.

In this paper, we attempt to test from first principles the assumption of local interactions. One of our main results is that the heuristic scaling $\delta v \sim k_\perp v^2 / k_\parallel V_A$ for three-wave interactions is questionable, and this casts serious doubts on the validity of the arguments given above of the energy spectrum $E(k_\perp) \propto \varepsilon^{1/2} k_\perp^{-2}$. We obtain a new nonlocal scaling for the fluctuation δv and explore its possible consequences for the anisotropic energy spectrum of weak MHD turbulence. Following Iroshnikov[2] and Kraichnan[3] we assume that three-wave interactions dominate, but our emphasis here is on the scaling of the anisotropy, neglected in the IK theory.

III. THREE-WAVE INTERACTIONS

The dominance of three-wave interactions in the limit of weak turbulence has been the subject of two recent papers by Montgomery and Matthaeus[18] and Ng and Bhattacharjee (NB).[14] Using the ideal MHD equations, NB show that three-wave interactions mediating the collision of two shear-Alfvén wave packets are in general nonzero if the $k_\parallel = 0$ components of the wave packets are nonzero. This has been shown to be true for the full as well as the reduced MHD (RMHD) equations.[19] Furthermore, using the RMHD equations, NB calculate in closed form the three-wave and four-wave interaction terms, and shown the former to be asymptotically dominant if the wave packets have nonzero $k_\parallel = 0$ components. To keep this discussion self-contained, we begin with a summary of relevant results by NB. The RMHD equations (in the ideal limit) can be written as

$$\frac{\partial \Omega}{\partial t} - \frac{\partial J}{\partial z} = [A, J] - [\phi, \Omega], \quad \frac{\partial A}{\partial t} - \frac{\partial \phi}{\partial z} = -[\phi, A],$$

where the magnetic field is given by $\mathbf{B} = \hat{\mathbf{z}} + \nabla_\perp A \times \hat{\mathbf{z}}$ with A as the magnetic flux function, the flow velocity is given by $\mathbf{v} = \nabla_\perp \phi \times \hat{\mathbf{z}}$ with ϕ as the stream function, and $[\phi, A] \equiv \phi_y A_x - \phi_x A_y$. The parallel vorticity is then $\Omega = -\nabla_\perp^2 \phi$, and the parallel current density is $J = -\nabla_\perp^2 A$. Note that we have normalized the background uniform magnetic field in the $\hat{\mathbf{z}}$ direction to have unit magnitude, and the density has been chosen so that the Alfvén speed $V_A = 1$. For weak interactions between two colliding shear-Alfvén wave packets $f^\pm$ traveling in the $\pm\hat{\mathbf{z}}$ directions, we write perturbative solutions of the form

$$\phi = f^-(\mathbf{x}_\perp, z^-) + f^+(\mathbf{x}_\perp, z^+) + \phi_1 + \phi_2 + \cdots,$$

$$A = f^-(\mathbf{x}_\perp, z^-) - f^+(\mathbf{x}_\perp, z^+) + A_1 + A_2 + \cdots,$$

where $\mathbf{x}_\perp = (x, y)$ is perpendicular to $\hat{\mathbf{z}}$ and $z^\pm = z \mp t$. For given zeroth order fields $f^\pm$, the first-order solutions can be found by solving the equations[14]

$$\frac{\partial \Omega_1}{\partial t} - \frac{\partial J_1}{\partial z} = 2\{[f^+, \nabla_\perp^2 f^-] + [f^-, \nabla_\perp^2 f^+]\} \equiv F,$$

$$\frac{\partial A_1}{\partial t} - \frac{\partial \phi_1}{\partial z} = 2[f^-, f^+] \equiv G. \tag{3}$$

This is a radiation equation for the first-order fields, with the source term determined by the overlap of the given zeroth-order fields f^+ and f^-. (The source term is localized both in space and in time, assuming that the functional forms of $f^\pm$

are chosen so that the wave packets are localized in z.) The asymptotic expression of ϕ_1, A_1 can be written as[14]

$$\phi_1(\mathbf{x}_\perp, t\to\infty)\to f_1^-(\mathbf{x}_\perp, z^-)+f_1^+(\mathbf{x}_\perp, z^+),$$

$$A_1(\mathbf{x}_\perp, t\to\infty)\to f_1^-(\mathbf{x}_\perp, z^-)-f_1^+(\mathbf{x}_\perp, z^+),$$

where

$$f_1^\pm(\mathbf{x}_\perp, z)=\pi\int [\tilde{F}'(\mathbf{k},\pm k_z)\mp\tilde{G}(\mathbf{k},\pm k_z)]e^{i\mathbf{kx}}\,d\mathbf{k},$$

with $\tilde{F}'(\mathbf{k},\omega)\equiv\tilde{F}(\mathbf{k},\omega)/k_\perp^2$ and $\tilde{F}(\mathbf{k},\omega)$ is the Fourier transform of $F(\mathbf{x},t)$, defined by

$$F(\mathbf{x},t)=\int \tilde{F}(\mathbf{k},\omega)e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)}\,d\mathbf{k}d\omega.$$

The Fourier transform $\tilde{G}(\mathbf{k},\omega)$ is similarly defined. To simplify the calculation that follows, we consider the case in which the functions $f^\pm(\mathbf{x}_\perp, z)$ are separable, i.e., $f^\pm(\mathbf{x}_\perp, z)=f_\perp^\pm(\mathbf{x}_\perp)f^\pm(z)$. [The calculation can also be carried through in the more general case when $f^\pm(\mathbf{x}_\perp, z)$ can be written as a sum of such separable terms, which is always possible as long as the boundary conditions in $\mathbf{x}_\perp$ are periodic.] Then we can write

$$F(\mathbf{x},t)=F_\perp(\mathbf{x}_\perp)f^+(z^+)f^-(z^-),$$

$$G(\mathbf{x},t)=G_\perp(\mathbf{x}_\perp)f^+(z^+)f^-(z^-),$$

$$\tilde{F}(\mathbf{k},\omega)=\frac{1}{2}\tilde{F}_\perp(\mathbf{k}_\perp)\tilde{f}^+(\kappa^+)\tilde{f}^-(\kappa^-),$$

$$\tilde{G}(\mathbf{k},\omega)=\frac{1}{2}\tilde{G}_\perp(\mathbf{k}_\perp)\tilde{f}^+(\kappa^+)\tilde{f}^-(\kappa^-),$$

where $\kappa^\pm\equiv(k_z\pm\omega)/2$, $\tilde{f}(\kappa^\pm)$ is the one-dimensional Fourier transform of $f^\pm(z^\pm)$, and $\tilde{F}_\perp$ and $\tilde{G}_\perp$ are the two-dimensional Fourier transforms of $F_\perp$ and $G_\perp$. We obtain

$$f_1^\pm(\mathbf{x}_\perp, z^\pm)=\pi u_\perp^\pm(\mathbf{x}_\perp)\tilde{f}^\pm(0)f^\pm(z^\pm)/2, \tag{4}$$

where

$$u_\perp^\pm(\mathbf{x}_\perp)=\int [\tilde{F}'_\perp\mp\tilde{G}_\perp]e^{i\mathbf{k}_\perp\cdot\mathbf{x}_\perp}\,d\mathbf{k}_\perp, \tag{5}$$

and $\tilde{f}^\pm(0)$ is the $k_z=0$ Fourier component of $f^\pm(z)$, with $\tilde{F}'(\mathbf{k}_\perp)\equiv\tilde{F}(\mathbf{k}_\perp)/k_\perp^2$. We see that the asymptotic (resonant) three-wave terms $f_1^\pm(\mathbf{x}_\perp, z^\pm)$ vanish in the special case $\tilde{f}^\mp(0)=0$, when the $k_z=0$ components of the wave packets are zero. Since this is a rather special case, we shall not consider this possibility further.

Note that $f_1^\pm$, given by Eq. (4), are bilinear functions of f^+ and f^-. Hence, in order to obtain the leading-order velocity fluctuation caused by three-wave interactions, f^+ and f^- need to be distinguished. This distinction is not made in the first estimate of Eq. (2) to which we shall return later.

IV. THREE-WAVE SPECTRUM: ANALYTICAL

We note that the expression (5) for three-wave interactions preserves the z dependence of the zeroth-order fields. This implies that there is no energy cascade parallel to the magnetic field. However, as pointed out by NB, four-wave interactions do not generally preserve the z dependence of the zeroth-order fields and exhibit harmonic generation, and can, in principle, contribute to a parallel cascade. Since three-wave couplings are much larger than four-wave couplings for weak turbulence, we neglect the effect of a parallel energy cascade. Using Eq. (5), we can then calculate explicitly the scaling of three-wave interactions. Specifically, our objective is to calculate the spectral indices of the three-wave fields as functions of the spectral indices of the zeroth-order fields. These functional relations should not be affected by the magnitude of the zeroth-order fields as long as we stay within the domain of validity of the weak turbulence approximation.

Imposing periodic boundary condition in $\mathbf{x}_\perp$, we can write

$$f_\perp^\pm(\mathbf{x}_\perp)=\sum_{mn} f_{mn}^\pm e^{2\pi i(mx+ny)},$$

where $f_{mn}^\pm$ are constants which define the spectral indices $\mu_\pm$ of the zeroth-order fields. We define the energy

$$\mathbf{E}_\pm=\int (\nabla_\perp f_\perp^\pm)^2\,d\mathbf{x}_\perp=\int E_\pm(k_\perp)dk_\perp, \tag{6}$$

with the spectral functions given by

$$E_\pm(k_\perp)\propto k_\perp^{-\mu_\pm} \text{ or } |f_{mn}^\pm|\propto(m^2+n^2)^{-(3+\mu_\pm)/4}.$$

Assuming that energy is distributed randomly in the zeroth-order fields according to these spectra, we now proceed to calculate the spectra of the first-order fields, as specified by Eq. (4). We rewrite the right-hand side of Eq. (5) as

$$u_\perp^\pm(\mathbf{x}_\perp)=\sum_{mn}[F'_{mn}\mp G_{mn}]e^{2\pi i(mx+ny)}$$

$$=\sum_{mn} u_{mn}^\pm e^{2\pi i(mx+ny)}, \tag{7}$$

with $F'_{mn}=F_{mn}/(2\pi)^2(m^2+n^2)$, and note from Eq. (3) that

$$F_\perp(\mathbf{x}_\perp)=2\{[f_\perp^+,\nabla_\perp^2 f_\perp^-]+[f_\perp^-,\nabla_\perp^2 f^+]\}$$

$$=\sum_{mn} F_{mn}e^{2\pi i(mx+ny)},$$

$$G_\perp(\mathbf{x}_\perp)=2[f_\perp^-, f_\perp^+]=\sum_{mn} G_{mn}e^{2\pi i(mx+ny)}.$$

By direct substitution, we obtain

$$G_{mn}=2(2\pi)^2\sum_{pq} f_{p,q}^+ f_{m-p,n-q}^-(qm-pn) \tag{8a}$$

$$=2(2\pi)^2\sum_{pq} f_{p,q}^- f_{m-p,n-q}^+(pn-qm), \tag{8b}$$

and

$$F'_{mn}=4(2\pi)^2\sum_{pq} f_{p,q}^+ f_{m-p,n-q}^-\frac{pm+qn}{m^2+n^2}(pn-qm)+G_{mn} \tag{9a}$$

$$=4(2\pi)^2\sum_{pq} f^-_{p,q}f^+_{m-p,n-q}\frac{pm+qn}{m^2+n^2}(pn-qm)-G_{mn}. \tag{9b}$$

Note that F'_{mn} is symmetric with respect to the exchange of the + and − indices, while G_{mn} is antisymmetric. Combining Eqs. (7)–(9), we obtain

$$u^{\pm}_{mn}=4(2\pi)^2\sum_{pq} f^{\pm}_{p,q}f^{\mp}_{m-p,n-q}\frac{pm+qn}{m^2+n^2}(pn-qm). \tag{10}$$

Now let the spectral indexes of $u^{\pm}_{mn}$ be $\nu_{\pm}$, i.e.,

$$|u^{\pm}_{mn}|\propto(m^2+n^2)^{-(3+\nu_{\pm})/4}.$$

We can estimate $\nu_{\pm}$ in two asymptotic regimes. For $\mu_+\gg\mu_-$, since f^+_{mn} decays much faster than f^-_{mn} at large wave numbers, we only need to sum over the domain of small (p,q) in Eqs. (8a) and (9a) for large (m,n). For $(m,n)\gg(p,q)$, we thus obtain

$$G_{mn}\approx 2(2\pi)^2 f^-_{mn}\sum_{pq} f^+_{p,q}(qm-pn)$$

$$=2(2\pi)^2 f^-_{mn}(\bar{q}m-\bar{p}n),$$

$$u^+_{mn}\approx 4(2\pi)^2 f^-_{mn}\sum_{pq} f^+_{p,q}\frac{pm+qn}{m^2+n^2}(pn-qm)$$

$$=4(2\pi)^2 f^-_{mn}\frac{\overline{pq}(n^2-m^2)+(\overline{p^2}-\overline{q^2})mn}{m^2+n^2},$$

where the overbar indicates average over the f^+_{mn} spectrum. It follows by inspection that

$$\nu_+\approx\mu_-,\quad \nu_-\approx\mu_--2. \tag{11a}$$

Similarly for $\mu_+\ll\mu_-$, we obtain

$$\nu_+\approx\mu_+-2,\quad \nu_-\approx\mu_+. \tag{11b}$$

V. THREE-WAVE SPECTRUM: NUMERICAL

The analytical estimates (11), although suggestive, cannot be assumed to hold in the regime $\mu_-\approx\mu_+$ which is of considerable physical interest. Therefore, we treat this regime numerically, as follows.

We calculate u^+_{mn} by a pseudo-spectral method. We first assign random amplitudes $f^{\pm}_{mn}$ for a given choice of zeroth-order spectral indices $\mu_{\pm}$. We then use fast Fourier transforms (FFT) to transform the data from spectral space to physical space. The nonlinear terms in physical space are evaluated by a finite-difference method. Finally, we find u^+_{mn} by another FFT, and calculate the spectral index ν_+ by either measuring the slope of a log–log plot of the spectrum, or by using a least-squares fit.

We remark that the numerical procedure discussed above is rather different from the procedure generally followed in incompressible MHD simulations that integrate forward in time the MHD equations of motion, and attempt to calculate the spectra, when possible, from a numerical identification of an inertial range. Our approach is based on our rigorous analytical formulas for three-wave interactions,[14]

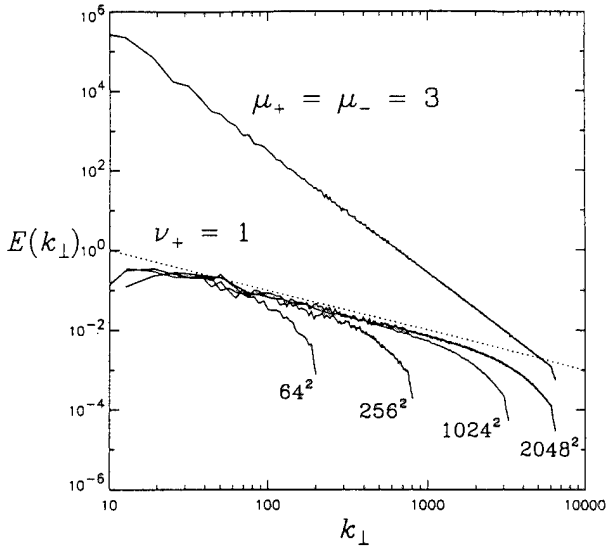

FIG. 1. The spectra of three-wave interactions calculated using different numerical resolution, with spectral index $\nu_+\approx 1$ and $\mu_+=\mu_-=3$. The spectra of the zeroth-order fields are also shown (arbitrary units).

valid in the limit of weak turbulence, which we use to obtain the relation between the spectral indices. In doing so, we assume that the spectral forms of the zeroth-order fields are given, and infer the indices of the first-order fields from a numerical study based on our analytical formulas.

In Fig. 1, we plot the spectra of the zeroth-order field $f^{\pm}_{mn}$ as well as the spectra of the first-order field u^+_{mn} for increasing spatial resolution (from 64^2 to 2048^2) for the case $\mu_+=\mu_-=3$. While the spectrum of the zeroth-order field is approximated very well by a straight line for almost the whole range of $k_\perp$, the spectrum of the first-order field curves downward for large $k_\perp$. Note, however, that as the spatial resolution is increased, the straight-line fit for the first-order spectra extends to larger values of $k_\perp$, with $\nu_+\approx 1$ for this case. Hence, it is reasonable to attribute the downward curve of the first-order spectra to the cutoff of the zeroth-order spectrum. As this cutoff is pushed to lower limits, the fit approaches a straight line to a higher degree of accuracy.

For 2048^2 resolution, Fig. 2 shows a contour plot of ν_+ in the range of $0\leq\mu_{\pm}\leq 5$. (We have also obtained, at lower resolution, numerical results in the larger range $-3\leq\mu_{\pm}\leq 20$, but the spectra are less clean and not shown here.) Note that the contour of ν_- is just that of ν_+ with the labels of $\mu_{\pm}$ interchanged, since $\nu_-(\mu_+,\mu_-)=\nu_+(\mu_-,\mu_+)$. The spectrum of G_{mn} has also been calculated and is symmetric with respect to $\mu_{\pm}$. It is found that the index of G_{mn} is given approximately by $\min(\mu_-,\mu_+)-2$ and

$$\nu_+\sim\min(\mu_-+2,\mu_+)-2,\quad \nu_-\sim\min(\mu_-,\mu_++2)-2, \tag{12a}$$

which is consistent with the asymptotic estimates (11). Moreover, for the case $\mu_+=\mu_-\equiv\mu$, we obtain

$$\nu\equiv\nu_{\pm}\sim\mu_{\pm}-2=\mu-2. \tag{12b}$$

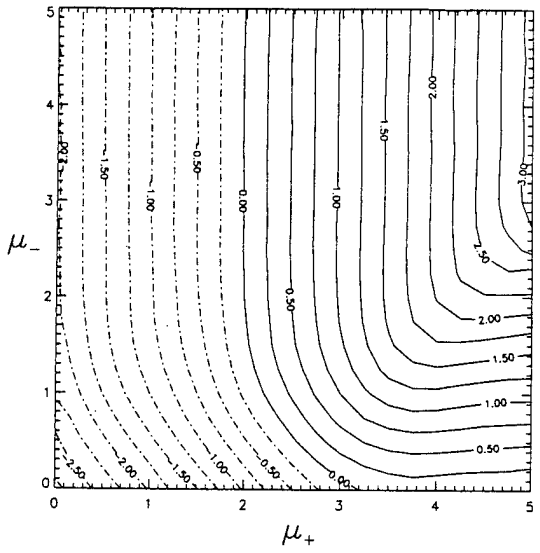

FIG. 2. Contour plots of the spectral index $\nu_\pm$ calculated using a resolution of 2048^2 within the range $0 \leq \mu_\pm \leq 5$.

Note that while the spectral indices of both F'_{mn} and G_{mn} are symmetric with respect to μ_+ and μ_-, that is not so for ν_+ or ν_-. This asymmetry is due to a significant cancellation between F'_{mn} and G_{mn} in the second term of Eq. (7).

VI. DISCUSSIONS

The results reported above are quite different from those expected by scaling arguments in Sec. II, which begin from the relation[4,5]

$$\frac{\delta v_k}{v_k} \sim \frac{k_\perp v_k}{k_\parallel V_A}. \tag{13}$$

Equation (13) is the first relation in Eq. (2), rewritten with a subscript k on δv as well as v in order to emphasize the assumption of local interactions. According to this assumption, the dominant contribution to δv_k comes from v_k which is an eddy of the same spatial scale (k^{-1}), so that $\nu \sim 2\mu - 3$. In contrast, the present calculation suggests that the scaling of the magnitude of the forward-propagating first-order field δv_k^+ should be written

$$\delta v_k^+ \sim \frac{k_\perp v_k^+ v^-}{k_\parallel V_A}. \tag{14}$$

where v^- represents the average amplitude of the backward-propagating wave packet which includes the effect of all perpendicular length scales (implying, in effect, nonlocal interactions in k-space). We emphasize that zeroth-order fields with typical magnitudes v^+ (propagating forward) and v^- (propagating backward) are uncorrelated since they are assumed to be Alfvén wave packets coming in from two opposite directions with different regions of origination. Therefore, we need to distinguish between v^- and v^+ in a way that the heuristic scaling (13) does not. This point can also be appreciated by inspecting Eqs. (4) and (5), which are consistent with the scaling (14). If we regard v^- in (14) as a constant independent of $k_\perp$, we recover the spectral index ν_+ given by Eq. (12b). Similar considerations also hold for ν_-.

On the other hand, if we impose the *ad hoc* requirement that v^- and v^+ are dominantly correlated only for the same length scale k^{-1}, then the summation in Eq. (10) will mainly pick up those local interaction terms so that

$$u_{mn}^\pm \sim f_{mn}^+ f_{mn}^- (m^2+n^2)^{3/2}.$$

[Note that the number of modes included in the summation is now proportional to $(m^2+n^2)^{1/2}$.] We then have

$$\nu^+ \sim \nu^- \sim \mu^+ + \mu^- - 3, \tag{15}$$

which is the same as the estimate $\nu \sim 2\mu - 3$ obtained from Eq. (13) for the case $\mu^+ = \mu^-$.

These subtle differences in scaling—obtained by imposing and relaxing the assumption of local interactions—have significant implications for the spectral index of the anisotropic energy spectrum. Let us explore some of these consequences by means of heuristic arguments. Following the procedure in Sec. II, let N_k be the number of collisions needed to distort the wave amplitude v_k with typical perpendicular wave number $k_\perp$. Then $N_k \sim (v_k/\delta v_k)^2 \sim E_k/E_{1k}$, where E_{1k} is the energy spectral function of the first-order field due to three-wave interactions. Assuming that there is only a perpendicular energy cascade, the energy transfer rate can be estimated to be

$$\varepsilon_k \sim \frac{v_k^2}{N_k \tau} \sim v_k^2 \left(\frac{\delta v_k}{v_k}\right)^2 k_\parallel V_A, \tag{16}$$

which should be constant in the inertial range, with $\tau \sim (k_\parallel V_A)^{-1}$. If we use $v_k^2 \sim E_k k_\perp$ and the estimate (13), we obtain the anisotropic IK spectrum $E_k \propto k_\perp^{-2}$. In contrast, if we use the estimate (14), we obtain $E_k \propto k_\perp^{-3}$. These results can also be obtained by an alternative argument, as follows. According to Eq. (16), we have $\varepsilon_k \propto k_\perp E_{1k}$ which suggests that the spectral indices $\mu_\pm$ should be chosen to correspond to $\nu_\pm = 1$. By looking up Fig. 2 (and the contour of ν_- which is, in fact, the same graph with the exchange of labels $\mu_+ \leftrightarrow \mu_-$), we obtain the result $\mu_\pm = 3$ for $\nu_\pm = 1$. This result should be contrasted with the result $\mu_\pm = 2$, or $\mu^+ + \mu^- = 4$, obtained from Eq. (15) by heuristic arguments based on the local interaction assumption, as well as with the result $\mu_\pm = \frac{3}{2}$ obtained by closure methods in the isotropic case.[20]

The new anisotropic spectrum, which is steeper than the anisotropic IK spectrum, can also be obtained by dimensional analysis. We write the energy cascade rate of the forward propagating wave in the form $\varepsilon^+ \sim k_\perp^\alpha \times E_k^{+\beta} (v^-)^2 (k_\parallel V_A)^{-1}$, where we assume that E_k^+ depends only on $k_\perp$ and that the wave number $k_\parallel$ appears always in the product with V_A as a parameter. We also use Eqs. (14) and (16) to obtain the dimensional dependence on v^-. Dimensional analysis then yields $\alpha = 3$, $\beta = 1$, and $\nu = 3$, which implies that $E_k \propto k_\perp^{-3}$.

We conclude this paper with a word of caution. The proposed scaling $E_k \propto k_\perp^{-3}$ is obtained here by heuristic arguments after Eq. (13) is replaced by the more rigorous scaling relation (14). While we have provided analytical and numerical justification of the scaling relation (14), we have not shown that the heuristic arguments used thereafter to deduce

the energy spectrum can be justified from first principles when the three-wave interactions are nonlocal. A critical evaluation of the validity of these commonly used heuristic arguments will require further analytical and numerical work and will be the subject of future research.

ACKNOWLEDGMENTS

We thank Professor P. Goldreich for several stimulating and instructive exchanges on MHD turbulence, and Dr. R. Kraichnan for providing us with a copy of his unpublished work (written in 1994), cited in this paper as Ref. 17.

This research is supported by the National Science Foundation Grant No. ATM 93-10157 and the Air Force Office of Scientific Research Grant No. F49620-93-10071. Numerical computations were carried out on the C90 at the San Diego Supercomputer Center.

[1]A. N. Kolmogorov, C. R. Akad. Sci. SSSR **30**, 301 (1941).
[2]P. S. Iroshnikov, Astron. Zh. **40**, 742 (1963).
[3]R. H. Kraichnan, Phys. Fluids **8**, 1385 (1965).
[4]A. Pouquet, P. L. Sulem, and M. Meneguzzi, Phys. Fluids **31**, 2635 (1988).
[5]D. Biskamp and H. Welter, Phys. Fluids B **1**, 1964 (1989).
[6]P. J. Coleman, Planet. Space Sci. **15**, 953 (1967).
[7]J. W. Belcher and L. Davis, J. Geophys. Res. **76**, 3534 (1971).
[8]L. F. Burlaga and J. M. Turner, J. Geophys. Res. **81**, 73 (1976).
[9]M. Dobrowolny, A. Mangeney, and P. Veltri, Phys. Rev. Lett. **45**, 144 (1980).
[10]D. Montgomery and L. Turner, Phys. Fluids **24**, 825 (1981).
[11]J. V. Shebalin, W. H. Matthaeus, and D. Montgomery, J. Plasma Phys. **29**, 525 (1983).
[12]S. Oughton, E. R. Priest, and W. H. Matthaeus, J. Fluid Mech. **280**, 95 (1994).
[13]S. Sridhar and P. Goldreich, Astrophys. J. **432**, 612 (1994).
[14]C. S. Ng and A. Bhattacharjee, Astrophys. J. **465**, 845 (1996).
[15]W. H. Matthaeus, S. Ghosh, S. Oughton, and D. A. Roberts, J. Geophys. Res. **101**, 7619 (1996).
[16]M. Hossain, P. C. Gray, D. H. Pontius, Jr., W. H. Matthaeus, and S. Oughton, Phys. Fluids **7**, 2886 (1995).
[17]S. Chen and R. H. Kraichnan, "Inhibition of turbulent cascade by sweep," submitted to J. Plasma Phys. (1996).
[18]D. Montgomery and W. H. Matthaeus, Astrophys. J. **447**, 706 (1995).
[19]H. R. Strauss, Phys. Fluids **19**, 134 (1976).
[20]R. Grappin, A. Pouquet, and J. Leorat, Astron. Astrophys. **126**, 51 (1983).

Finite-time vortex singularity and Kolmogorov spectrum in a symmetric three-dimensional spiral model

A. Bhattacharjee, C. S. Ng, and Xiaogang Wang
Department of Physics and Astronomy, University of Iowa, Iowa City, Iowa 52242
(Received 5 October 1994; revised manuscript received 30 May 1995)

A recent analytical model of three-dimensional Euler flows [Phys. Rev. Lett. **69**, 2196 (1992)] which exhibits a finite-time vortex singularity is developed further. The initial state is symmetric and contains a velocity null (stagnation point) which is collinear with two vorticity nulls. Under some assumptions, it is shown by asymptotic analysis of the Euler equation that the vorticity blows up at the stagnation point as inverse time in a locally self-similar manner. The spatial structure of the inviscid flow in the vicinity of the singularity involves disparate small scales. The effect of a small but finite viscosity is shown to arrest the formation of the singularity. The presence of spiral structure in the initial conditions leads naturally to the model developed by Lundgren [Phys. Fluids **25**, 2193 (1982)] in which the gradual tightening of spirals by differential rotation provides a mechanism for transfer of energy to small spatial scales. It is shown by asymptotic analysis of the Navier-Stokes equation, that a time-average over the lifetime of the spiral vortex in the present model yields the Kolmogorov spectrum.

PACS number(s): 47.27.Cn

I. INTRODUCTION

A fundamental problem in fluid dynamics is the derivation of the Kolmogorov spectrum for incompressible turbulent flows from the Navier-Stokes equation. The Kolmogorov scaling law [1], originally derived by dimensional analysis, has been validated by several experiments and numerical simulations, but its theoretical derivation from the underlying dynamical equations has remained a challenge for over 50 years.

It is widely believed that a first step in the development of a dynamical theory of turbulence should be the identification of a mechanism by which energy in the large spatial scales can be transferred to the small scales. This has motivated the search for finite-time vortex singularities in three-dimensional Euler flows, since in two-dimensional flows that tend to zero at infinity and evolve from smooth initial conditions, the formation of a finite-time vortex singularity is forbidden [2–4].

Theoretical models that may yield finite-time vortex singularities in three-dimensional Euler flows have been the subject of many investigations, most of which are numerical [5–9]. Until 1990, the results were inconclusive despite the sophistication of the numerical methods employed. Recently, finite-time singularities have been reported in axisymmetric flows with swirl [10,11], but doubts have been raised that the growth of vorticity observed in these studies may be exponential [12]. A related problem has been investigated by Childress [13], who has interpreted the occurrence of a finite-time singularity in a nearly two-dimensional flow as a signature of the loss of near-two-dimensionality. E and Shu [14] present a numerical study of two-dimensional Boussinesq convection (also studied earlier by Grabowski and Clark [15]) and report exponential growth of vorticity in those regions where Pumir and Siggia [11] find a finite-time singularity. Despite the depth of effort that has gone into these investigations, it cannot be said that the issue of vorticity intensification in axisymmetric flows with swirl has been settled beyond doubt. E and Shu find that the growth of vorticity is much more intense on the side of a rising bubble than the cap where the growth of vorticity is found to be exponential. Caflisch [16] has presented a different viewpoint on this problem and has demonstrated the development of finite-time vortex singularities in a model of complex-valued axisymmetric flows with swirl. However, the singularities found by Caflisch do not occur where they do in Refs. [10] and [11], but at the centers of the rolls. Though Caflisch's results are mathematically interesting, there are questions regarding the relevance of his findings to the real dynamics of the Euler equation.

In a way, the controversy surrounding the problem of axisymmetric flows with swirl encapsulates the difficulties involved in the search for finite-time vortex singularities. The question is not only *how rapidly in time* vorticity grows, but also *where* the singular growth of vorticity should occur. The first question involves dynamics, whereas the second involves geometry. In attempting to answer the second question, we adopt the point of view that vortex singularities will tend to occur near separatrices. We arrive at this point of view, prompted by insights on the analogous problem of current singularity formation in magnetohydrodynamics (MHD). Though there are profound differences between the dynamics of magnetic fields in plasmas and velocity fields in Euler flows, there are some geometrical similarities. We discuss below some of the similarities and differences.

Magnetic fields **B** are divergence-free, as are incompressible velocity fields **v**. The analogy between **B** and **v** is most obvious when we consider steady solutions of the Euler equation

$$\mathbf{v}\times\boldsymbol{\omega}=\nabla h\ ,\quad \nabla\times\mathbf{v}=\boldsymbol{\omega}\ , \tag{1}$$

where ω is the vorticity and $h=p/\rho+v^2/2$ is the Ber-

 ©1995 The American Physical Society

noulli function for a fluid of pressure p and density ρ. Equation (1) is analogous to the magnetostatic condition [17]

$$\mathbf{J}\times\mathbf{B}=\nabla p, \quad \nabla\times\mathbf{B}=\mathbf{J}\ , \tag{2}$$

where $\mathbf{J}$ is the electrical current density and p is the fluid pressure. A comparison of (1) and (2) suggests the analogy $\mathbf{B}\leftrightarrow\mathbf{v}$, $\mathbf{J}\leftrightarrow\boldsymbol{\omega}$, $p\leftrightarrow h_0-h$ (where h_0 is a constant). It is this analogy that has provided the basis for certain deductions regarding the properties of steady three-dimensional Euler flows in Refs. [18–22]. In Ref. [22], solutions of (1) that are topologically toroidal are considered and it is demonstrated that δ-function vortex singularities can occur at the so-called rational surfaces of a nearly integrable velocity field. These rational surfaces, on which the streamlines close on themselves, are the source of separatrices in a torus. The vortex singularities at the rational surfaces are exactly where current singularities occur in the toroidal solutions of a nearly integrable magnetic field [21].

This analogy between $\mathbf{B}$ and $\mathbf{v}$ in steady state breaks down when we consider dynamics. In an ideal plasma, the magnetic field $\mathbf{B}$ obeys the induction equation

$$\frac{\partial\mathbf{B}}{\partial t}-\nabla\times(\mathbf{v}\times\mathbf{B})=0\ , \tag{3}$$

whereas the vorticity $\boldsymbol{\omega}$ (and *not* the velocity) obeys the (analogous) Euler equation

$$\frac{\partial\boldsymbol{\omega}}{\partial t}-\nabla\times(\mathbf{v}\times\boldsymbol{\omega})=0\ . \tag{4}$$

In other words, for dynamical evolution, the appropriate analogy is $\mathbf{B}\leftrightarrow\boldsymbol{\omega}$ (to be contrasted with $\mathbf{J}\leftrightarrow\boldsymbol{\omega}$ in steady state). The ramifications of this dynamical analogy have been partially explored in Refs. [19–22] and has prompted Greene [23] to suggest "...that an essential aspect of turbulence is that it is a dynamo for vorticity amplification..." It is because of this analogy that mathematical methods used in MHD have proved to be valuable in obtaining new stability results in Euler flows with stagnation points [24,25]. The analogy suggests that since in MHD magnetic-field nulls are a possible source of separatrices [26,27] where singular currents tend to grow, the neighborhood of vorticity and velocity nulls are possible sites for singularity formation in three-dimensional Euler flows.

Recently, we have proposed an analytical model [28] of three-dimensional Euler flows containing nulls. The initial state of this model (discussed in Sec. II) is symmetric and has a stagnation point at the origin, flanked by two vorticity nulls. The straight line joining the two vorticity nulls intersects the stagnation point at the origin, which is preserved for all times by the flow. In Ref. [28] we have attempted to show that this model yields a finite-time singularity by means of a multiple-scale analysis under some strong assumptions regarding the form of the solutions. In Sec. III we give a modified version of this analysis that makes the derivation of the singularity more transparent. The modified analysis, which is carried out to higher order, demonstrates that there are two disparate spatial scales in the neighborhood of the stagnation point. As in Ref. [28], we find that the vorticity blows up as $(t_c-t)^{-1}$ at the stagnation point in a locally self-similar manner within a collapsing inner region. While the velocity remains bounded in the inner region, it blows up as $(t_c-t)^{-1/2}$ in a surrounding region where the flow violates the self-similar scaling of the inner region.

The results obtained in Sec. III by the perturbative analysis motivate a more general treatment of stagnation point flows (Sec. IV) in which the pressure is calculated self-consistently by a (nonlocal) integral relation involving the velocity. This generalized treatment enables us to consider a wider class of initial conditions (of finite energy) than that considered in Sec. III. The results of Sec. IV support the results obtained in Sec. III and furthermore suggest that there are broadly two types of vortex singularities that can be realized with initial conditions of high symmetry.

In Sec. V, we study the models discussed in Secs. III and IV in the complex spatial domain, following the recent treatment of Tanveer and Speziale [29]. In this approach, the Euler equation is continued into the complex unphysical domain. Since a class of smooth and bounded initial conditions in the real physical domain can be singular when analytically continued into the complex unphysical domain, the question of finite-time singularities in the real physical domain can be reduced to the question of whether the singularities in the complex domain reach the real domain finite time. The analytical framework developed in Ref. [29] is based on a crucial set of assumptions that have not been shown to hold in general. Whether these assumptions hold for the symmetric dynamics considered in this paper also remains unproved. However, the approach of Tanveer and Speziale does reproduce the two disparate small scales found by the asymptotic analysis of Sec. III.

We discuss in Sec. VI that the analytical results of Secs. III–V do not violate the rigorous constraints due to Constantin and Fefferman [30,31]. However, we point out that the relevant theorems in Refs. [30] and [31] are proved under a specific assumption that appears to be violated by self-similar flows and hence it is questionable whether these theorems are directly applicable to our model.

One of the interesting features of the smooth initial state in Ref. [28] is the presence of spiral structures. Though the spiral structure does not play a significant role in the inviscid evolution of the flow, its occurrence is reminiscent of the strained vortex model of Lundgren [32–34]. In Lundgren's model, nonaxisymmetric coherent vortices with spiral structures interact with each other only through the mediation of an axisymmetric, background straining flow. By means of detailed asymptotic analysis, Lundgren demonstrates that the Kolmogorov spectrum then follows from the Navier-Stokes equation.

Lundgren's discovery, and the natural presence of spirals in our model, motivates us to investigate the effect of viscosity (Sec. VII). It is shown that the presence of viscosity arrests the formation of the finite-time vortex

singularity in our model. As the vorticity intensifies at the velocity null, the tightening of the spirals in our initial state provides a mechanism for transfer of energy from the large to the small spatial scales. An asymptotic analysis based on the Navier-Stokes equation leads us quite naturally to Lundgren's considerations and we too obtain the Kolmogorov spectrum. Viewed in its totality, the present model possesses some features that enhance Lundgren's conception. In particular, the model describes the dynamics of a three-dimensional (but symmetric) flow that evolves from smooth initial conditions and tends to a finite-time singularity (which Lundgren's model does not). Subsequently, viscosity thwarts the formation of the singularity and yields the Kolmogorov spectrum (as Lundgren's model does). Thus the present model can claim to capture some features of decaying hydrodynamic turbulence.

II. INITIAL CONDITIONS

At $t=0$, we consider a symmetric flow $\mathbf{u}$ of the form

$$u_x=f(y)\ , \tag{5a}$$

$$u_y=f(z)\ , \tag{5b}$$

$$u_z=f(x)\ . \tag{5c}$$

For specificity, we choose

$$f(x)=u_0\frac{x}{a_0}\exp\left[-\frac{\varepsilon x^2}{a_0^2}\right]\ , \tag{6}$$

where u_0 and a_0 are positive constants and ε is a small and positive parameter (i.e., $\varepsilon\ll 1$) that separates the local scale a_0 from the larger scale of the globally extended Gaussian envelope of width $\varepsilon^{-1/2}a_0$. The origin is a stagnation point or a null of the velocity field. Near the origin, we have

$$\mathbf{u}=\mathbf{x}\cdot(\nabla\mathbf{u})_0\ , \tag{7}$$

where the tensor

$$(\nabla\mathbf{u})_0=f'(0)\begin{bmatrix}0&1&0\\0&0&1\\1&0&0\end{bmatrix} \tag{8}$$

has the eigenvalues $\lambda_0=f'(0)$ and $\lambda_\pm=f'(0)\exp(\pm i2\pi/3)$. According to the standard nomenclature [35,36] since $f'(0)=u_0/a_0>0$, this null is of type A_s. Near the null, the eigenvectors for the complex eigenvalues $\lambda_\pm$ lie on the Σ_{A_s} surface, which is the stable manifold, whereas the eigenvector for the real and positive eigenvalue λ_0 lies on the γ_{A_s} line, which is the unstable manifold. The subscript s denotes the spiraling trajectories of the streamlines into the null in the Σ_{A_s} surface.

From the relation $\boldsymbol{\omega}=\nabla\times\mathbf{u}$, we obtain

$$\omega_x=-f'(z)=-\frac{u_0}{a_0}\left[1-\frac{2\varepsilon z^2}{a_0^2}\right]\exp\left[-\frac{\varepsilon z^2}{a_0^2}\right]\ , \tag{9a}$$

$$\omega_y=-f'(x)=-\frac{u_0}{a_0}\left[1-\frac{2\varepsilon x^2}{a_0^2}\right]\exp\left[-\frac{\varepsilon x^2}{a_0^2}\right]\ , \tag{9b}$$

$$\omega_z=-f'(y)=-\frac{u_0}{a_0}\left[1-\frac{2\varepsilon y^2}{a_0^2}\right]\exp\left[-\frac{\varepsilon y^2}{a_0^2}\right]\ . \tag{9c}$$

At the origin, $\omega_x=\omega_y=\omega_z=-u_0/a_0$. It follows from Eqs. (9) that the vorticity has two nulls at $x=y=z=\pm a_0/\sqrt{2\varepsilon}\equiv a_\pm$. The vorticity null $x=y=z=a_+$ is of type A_s, whereas the null $x=y=z=a_-$ is of type B_s. A null of type B_s is characterized by one real, negative eigenvalue and two complex eigenvalues. The eigenvector for the real, negative eigenvalue lies on the γ_{B_s} curve, which is a stable manifold. The eigenvectors for the two complex eigenvalues lie on a two-dimensional plane that coincides with the Σ_{B_s} surface, which is an unstable manifold.

Expanding Eqs. (5), (6), and (9) in Taylor series, we obtain

$$u_x=s_0y\left[1-\frac{\varepsilon y^2}{a_0^2}\right]+O(\varepsilon^2)\ . \tag{10a}$$

$$u_y=s_0z\left[1-\frac{\varepsilon z^2}{a_0^2}\right]+O(\varepsilon^2)\ , \tag{10b}$$

$$u_z=s_0x\left[1-\frac{\varepsilon x^2}{a_0^2}\right]+O(\varepsilon^2)\ , \tag{10c}$$

whence

$$\omega_x=-s_0\left[1-\frac{3\varepsilon z^2}{a_0^2}\right]+O(\varepsilon^2)\ , \tag{11a}$$

$$\omega_y=-s_0\left[1-\frac{3\varepsilon x^2}{a_0^2}\right]+O(\varepsilon^2)\ , \tag{11b}$$

$$\omega_z=-s_0\left[1-\frac{3\varepsilon y^2}{a_0^2}\right]+O(\varepsilon^2)\ , \tag{11c}$$

where $s_0=u_0/a_0$. The symmetry of the initial conditions singles out the $x=y=z$ line as a natural axis. We introduce a new coordinate system (x',y',z'), where $\hat{\mathbf{z}}'=(1/\sqrt{3})(1,1,1)$ and $\hat{\mathbf{x}}',\hat{\mathbf{y}}'$ are two mutually orthogonal unit vectors in the plane normal to $\hat{\mathbf{z}}'$, with $\hat{\mathbf{x}}'=(1/\sqrt{6})(-1,-1,2)$ and $\hat{\mathbf{y}}'=(1/\sqrt{2})(1,-1,0)$. The initial flow (10) can be written as

$$\mathbf{u}=\begin{bmatrix}u_{x'}\\u_{y'}\\u_{z'}\end{bmatrix}=s_0\begin{bmatrix}-1/2&\sqrt{3}/2&0\\-\sqrt{3}/2&-1/2&0\\0&0&1\end{bmatrix}\begin{bmatrix}x'\\y'\\z'\end{bmatrix}+O(\varepsilon)\ . \tag{12}$$

Defining $x'=r'\cos(\theta'+\pi/6)$, $y'=r'\sin(\theta'+\pi/6)$, Eq. (12) can be transformed to

$$\mathbf{u}=\begin{bmatrix}u_r\\u_\theta\\u_z\end{bmatrix}=s_0\begin{bmatrix}-r/2\\-(\sqrt{3}/2)r\\z\end{bmatrix}+O(\varepsilon)\ , \tag{13}$$

where we have dropped the primes for notational convenience. As noted in Ref. [28], Eq. (13) is axisymmetric to leading order. The departure from axisymmetry and the spiral structure is manifest at higher order. The initial vorticity is given by the expansion

$$\boldsymbol{\omega}=-\sqrt{3}s_0\left[\begin{bmatrix}0\\0\\1\end{bmatrix}+\frac{\varepsilon z}{a_0^2}\begin{bmatrix}r\\-\sqrt{3}r\\-z\end{bmatrix}-\frac{\varepsilon r^2}{\sqrt{2}a_0^2}\begin{bmatrix}\sin 3\theta\\\cos 3\theta\\\sqrt{2}\end{bmatrix}\right]+O(\varepsilon^2)\ . \tag{14}$$

In what follows, we shall investigate the time evolution of the initial state described above according to the Euler equation

$$\frac{\partial\boldsymbol{\omega}}{\partial t}+\mathbf{u}\cdot\nabla\boldsymbol{\omega}=\boldsymbol{\omega}\cdot\nabla\mathbf{u}\ . \tag{15}$$

Here we have done away with the so-called background flow considered in Ref. [28], which is inessential for singularity formation.

III. PERTURBATIVE SOLUTIONS

The localized flow $\mathbf{u}$ obeys the symmetry relations

$$u_x(x,y,z)=u_y(z,x,y)=u_z(y,z,x)\ . \tag{16}$$

Since these symmetry relations are assumed to hold at $t=0$, they hold for all times [37]. We consider solutions of the form

$$u_x(\mathbf{x},t)=s_2(\mathbf{x},t)y\exp\left[-\frac{y^2}{a_2^2(\mathbf{x},t)}\right]\ , \tag{17a}$$

$$u_y(\mathbf{x},t)=s_3(\mathbf{x},t)z\exp\left[-\frac{z^2}{a_3^2(\mathbf{x},t)}\right]\ , \tag{17b}$$

$$u_z(\mathbf{x},t)=s_1(\mathbf{x},t)x\exp\left[-\frac{x^2}{a_1^2(\mathbf{x},t)}\right]\ . \tag{17c}$$

The six functions on the right-hand side of (17) are constrained by the incompressibility condition and the (two) symmetry relations (16). Exploiting the presence of the small parameter ε in the initial conditions, we seek solutions of the form

$$a_i=a(t)+\varepsilon\tilde{a}_i(\mathbf{x},t)\ , \tag{18a}$$

$$s_i=s(t)+\varepsilon\tilde{s}_i(\mathbf{x},t)\ , \tag{18b}$$

where $a(0)=a_0$ and $s(0)=s_0$. The form of Eq. (18) is based on the assumption that the solutions are locally self-similar. The solutions are not globally self-similar; the space dependence of the higher-order terms in Eq. (18) preclude such a possibility.

For the initial conditions discussed above, we develop solutions of the Euler equation by a formal perturbation expansion. We write

$$u_r=u_r^{(0)}+\varepsilon u_r^{(1)}+O(\varepsilon^2)\ , \tag{19a}$$

$$u_\theta=u_\theta^{(0)}+\varepsilon u_\theta^{(1)}+O(\varepsilon^2)\ , \tag{19b}$$

$$u_z=u_z^{(0)}+\varepsilon u_z^{(1)}+O(\varepsilon^2)\ , \tag{19c}$$

and

$$\omega_r=\omega_r^{(0)}+\varepsilon\omega_r^{(1)}+O(\varepsilon^2)\ , \tag{20a}$$

$$\omega_\theta=\omega_\theta^{(0)}+\varepsilon\omega_\theta^{(1)}+O(\varepsilon^2)\ , \tag{20b}$$

$$\omega_z=\omega_z^{(0)}+\varepsilon\omega_z^{(1)}+O(\varepsilon^2)\ . \tag{20c}$$

Using the symmetry relation (16) and the assumption of local self-similarity, the leading-order solution for the velocity can be written in the form

$$\mathbf{u}^{(0)}=\begin{bmatrix}u_r^{(0)}\\u_\theta^{(0)}\\u_z^{(0)}\end{bmatrix}=s(t)\begin{bmatrix}-r/2\\-(\sqrt{3}/2)r\\z\end{bmatrix}\ . \tag{21}$$

Then the z component of the leading-order Euler equation (15) for $\omega_z^{(0)}=-\sqrt{3}s(t)$ yields

$$\frac{\partial\omega_z^{(0)}}{\partial t}=s\omega_z^{(0)} \tag{22}$$

or

$$\frac{\partial s}{\partial t}=s^2\ . \tag{23}$$

Equation (23) has the exact solution

$$s=\frac{1}{t_c-t}\ , \tag{24}$$

which yields a finite-time singularity at $t_c\equiv 1/s_0=a_0/u_0$. The singularity occurs in a locally self-similar manner in the neighborhood of the velocity null. For all times prior to the blowup of vorticity, there exists a small region surrounding the velocity null where the solution is invariant under the scaling $\mathbf{x}\to c\mathbf{x}$, $\mathbf{u}\to\mathbf{u}, t\to ct$, where c is a constant. In this (shrinking) region, the solutions $u_x=y/(t_c-t)$, $u_y=z/(t_c-t)$, and $u_z=x/(t_c-t)$, which yield $\omega_x=\omega_y=\omega_z=-(t_c-t)^{-1}$, satisfy the Euler equation (15) exactly.

It is useful to calculate the pressure Hessian $\pi_{ij}\equiv\partial^2 p/\partial x_i\partial x_j$ for the leading-order solution. We obtain

$$\pi_{ij}^{(0)}=-\begin{bmatrix}0&s^2&s^2\\s^2&0&s^2\\s^2&s^2&0\end{bmatrix}\ . \tag{25}$$

Whereas the diagonal elements, which represents the local contributions to the pressure Hessian, are zero, the off-diagonal elements, which represent the nonlocal contributions, are nonzero (and eventually blow up). Hence the present model cannot be described by the so-called "restricted Euler system" of Leorat [38], Viellefosse [39], and Cantwell [40], who assume that

$\pi_{ij}=(\partial^2 p/\partial x_k \partial x_k)(\delta_{ij}/3)$.

We now consider the first-order solutions. To this end, we note that Eqs. (21) and (24) yield the leading-order Lagrangian equations

$$\frac{dr}{dt}=u_r^{(0)}=-\frac{r}{2(t_c-t)}\ , \tag{26a}$$

$$\frac{d\theta}{dt}=\frac{u_\theta^{(0)}}{r}=-\frac{\sqrt{3}}{2}\frac{1}{t_c-t}\ , \tag{26b}$$

$$\frac{dz}{dt}=u_z^{(0)}=\frac{z}{t_c-t}\ , \tag{26c}$$

which can be integrated to give

$$r=r_0\left[\frac{t_c-t}{t_c}\right]^{1/2}\ , \tag{27a}$$

$$\theta=\theta_0-\frac{\sqrt{3}}{2}\int_0^t dt' s(t')\ , \tag{27b}$$

and

$$z=z_0\frac{t_c}{t_c-t}\ . \tag{27c}$$

respectively.

Writing $\omega=\omega^{(0)}+\varepsilon\omega^{(1)}+\cdots$, we obtain the ($z$ component of the) first-order Euler equation

$$\frac{\partial\omega_z^{(1)}}{\partial t}+(\mathbf{u}\cdot\nabla\omega_z)^{(1)}=(\omega\cdot\nabla u_z)^{(1)}\ . \tag{28}$$

Motivated by Lundgren [32], we introduce the variable transformations

$$\omega_z^{(1)}(r,\theta,z,t)=S(t)\Omega(\xi,\vartheta,v,T)\ , \tag{29a}$$

$$u_z^{(1)}(r,\theta,z,t)=S^{1/2}(t)\Lambda(\xi,\vartheta,v,T)\ , \tag{29b}$$

where

$$S(t)=\exp\left[\int_0^t dt' s(t')\right] \tag{30}$$

is the total stretching in time t and the variable ξ, ϑ, v, and T are defined, respectively, as

$$\xi\equiv S^{1/2}r\ , \tag{31a}$$

$$\vartheta\equiv\theta+\frac{\sqrt{3}}{2}\int_0^t dt' s(t')\ , \tag{31b}$$

$$v\equiv z/S\ , \tag{31c}$$

and

$$T\equiv\int_0^t dt' S(t')\ . \tag{31d}$$

We now calculate the various terms in (28):

$$\left[\frac{\partial\omega_z}{\partial t}\right]^{(1)}=\frac{\partial\omega_z^{(1)}}{\partial t}=\frac{\partial S}{\partial t}\Omega+S\frac{\partial\Omega}{\partial t}=sS\Omega+S\frac{\partial\Omega}{\partial t}=s\omega_z^{(1)}+sS\frac{\partial\Omega}{\partial T}+\frac{s\xi}{2}\left[\frac{\partial\omega_z^{(1)}}{\partial\xi}+\sqrt{3}\frac{\partial\omega_z^{(1)}}{\xi\partial\vartheta}\right]-sv\frac{\partial\omega_z^{(1)}}{\partial v}\ , \tag{32a}$$

$$(\mathbf{u}\cdot\nabla\omega_z)^{(1)}=\mathbf{u}^{(1)}\cdot\nabla\omega_z^{(0)}+\mathbf{u}^{(0)}\cdot\nabla\omega_z^{(1)}\cong\mathbf{u}^{(0)}\cdot\nabla\omega_z^{(1)}=-\frac{s\xi}{2}\left[\frac{\partial\omega_z^{(1)}}{\partial\xi}+\sqrt{3}\frac{\partial\omega_z^{(1)}}{\partial\vartheta}\right]+sv\frac{\partial\omega_z^{(1)}}{\partial v}\ , \tag{32b}$$

$$(\omega\cdot\nabla u_z)^{(1)}=\omega^{(1)}\cdot\nabla u_z^{(0)}+\omega^{(0)}\cdot\nabla_z^{(1)}=s\omega_z^{(1)}-\sqrt{3}s\frac{\partial u_z^{(1)}}{\partial z}=s\omega_z^{(1)}-\sqrt{3}sS^{-1/2}\frac{\partial\Lambda}{\partial v}\ . \tag{32c}$$

Using Eqs. (32), the first-order Euler equation (28) can be reduced to

$$\frac{\partial\Omega}{\partial T}=-\sqrt{3}\frac{1}{S^{3/2}}\frac{\partial\Lambda}{\partial v}\ . \tag{33}$$

As $t\to t_c$, the term on the right-hand side of (33) can be neglected. Hence, in this limit

$$\frac{\partial\Omega}{\partial T}=0\ , \tag{34}$$

which implies that

$$\Omega=\Omega(\xi,\vartheta,v)=\Omega_0(\xi_0(\xi,\theta,v),\vartheta_0(\xi,\theta,v),v_0(\xi,\theta,v))\ , \tag{35}$$

where

$$\Omega_0=\omega_z^{(1)}(r,\theta,z,0)=\tilde{\omega}(r,\theta,z)=\sqrt{3}s_0\left[\frac{r^2+z^2}{a_0^2}\right] \tag{36}$$

is the initial condition for the higher-order solution. Using (29a), we can write

$$\begin{aligned}\omega_z^{(1)}(r,\theta,z;t)&\cong S(t)\tilde{\omega}\left[r\left[\frac{t_c}{t_c-t}\right]^{1/2},\theta+\frac{\sqrt{3}}{2}\int_0^t dt' s(t'),z\left[\frac{t_c-t}{t_c}\right]\right]\\&=\frac{\sqrt{3}s_0}{a_0^2}\exp\left[\int_0^t s(t')dt'\right]\left[r^2\left[\frac{t_c}{t_c-t}\right]+z^2\left[\frac{t_c-t}{t_c}\right]^2\right]\ .\end{aligned} \tag{37}$$

If follows by inspection of Eq. (37) that the first-order solution has a small radial scale, collapsing as $(t_c-t)^{1/2}$, while it stretches axially along z. The collapsing small radial scale of the first-order solution is thus different from the small scale of the leading-order solution, collapsing as (t_c-t).

The structure of the small scale obtained above can be understood simply by invoking the law of conservation of vorticity strength [41]. We consider a "long-thin" cylindrical vortex tube of strength $\omega_z A=\text{const}$, where $A=\pi r^2$. Since $\omega_z \propto (t_c-t)^{-1}$, it follows that $r \propto (t_c-t)^{1/2}$. While the velocity is bounded in the inner region, which collapses as (t_c-t), the Lagrangian equations (26) imply that the velocity must blow up on the "larger" small scale, which collapses as $(t_c-t)^{1/2}$.

The analysis given above is perturbative and we have not obtained closed-form solutions. The calculation has been essentially carried through the first two orders, which make it apparent that more than one small spatial scale is involved in this inviscid problem. It is possible to extend the calculation, in principle, to higher orders, but the equations are complicated and do not seem amenable to an analytical solution.

A pertinent question is whether the singularity obtained in our model is unphysical since the initial flows have infinite total energy. As discussed in Ref. [42], this is indeed a limitation of earlier investigations of two [42] and three-dimensional [43] solutions of the stagnation-point form. We emphasize an important difference between our system of flows and those considered in Refs. [42] and [43]: in our initial conditions, the velocity (and vorticity fields) are bounded everywhere, including points at infinity. This means that the energy density is initially finite everywhere, including points at infinity. Infinite energy is obtained in our initial conditions merely because our system size is infinite, but the finite-time vortex singularity is not an artifact of the infinite system size.

It is well known that the Euler equation is an integro-differential system. (For instance, the velocity must be calculated self-consistently from the vorticity by carrying out an integration over the whole space with suitable boundary conditions). In Ref. [28] a lengthy analysis (which will not be repeated here) is given to match "local" flows of the form (17) to a "global" symmetric flow of general functional form. We assume here that such global solutions do exist (that is, they satisfy the integro-differential system). To further strengthen this conclusion, we give a generalized analysis in the next section that treats the integro-differential system explicitly.

IV. GENERALIZATIONS OF THE MODEL

The model discussed above can be generalized by keeping its basic symmetric features, but allowing for a more general form of the solution. We write

$$\mathbf{u}(\mathbf{x},t)=[u(x,y,z,t),u(y,z,x,t),u(z,x,y,t)] \tag{38a}$$

and assume that in the vicinity of the origin the Taylor expansion of the flow exists and that it only has odd-order terms, i.e.,

$$u(x,y,z,t)=\sum_{l,m,n\geq 0} a_{lmn}(t)x^l y^m z^n , \tag{38b}$$

where $a_{lmn}(t)=0$ if $l+m+n$ is even. [Specifically, it is reasonable to assume that the Taylor series is convergent for $|\mathbf{x}|<\delta(t)$, where δ is finite and we allow $\delta(t\rightarrow t_c)\rightarrow 0$.] The symmetry relations (38a) ensure that we have a velocity stagnation point at the origin for all time since the Euler equation preserves these relations, if satisfied initially. We also assume that $u(|\mathbf{x}|\rightarrow\infty, t=0)\rightarrow 0$ sufficiently rapidly that the initial flow has finite energy. Then flows with bounded energy can be constructed by treating the Euler equation as the integro-differential system

$$\frac{\partial \mathbf{u}}{\partial t}+\mathbf{u}\cdot\nabla\mathbf{u}=-\nabla p , \tag{39a}$$

$$p(\mathbf{x},t)=\frac{1}{4\pi}\int \frac{\nabla\cdot(\mathbf{u}\cdot\nabla\mathbf{u})}{|\mathbf{x}-\mathbf{x}'|}d\mathbf{x}' , \tag{39b}$$

where $\rho=1$ and the condition $\nabla\cdot\mathbf{u}=0$ is satisfied, once it is so initially. It is easy to see that the pressure p is still finite at large $\mathbf{x}$ even if $\mathbf{u}$ has a locally self-similar singularity, because such a singularity occurs only in a volume that also tends to zero. This ensures that the flow will remain vanishingly small for large $\mathbf{x}$ and thus the total energy will remain finite for all time, since it is so initially.

We consider the general form of the first order expansion of u,

$$u(x,y,z,t)=b(t)y+c(t)z+O(|\mathbf{x}|^3) , \tag{40}$$

which is consistent with the condition $\nabla\cdot\mathbf{u}=0$. The vorticity $\boldsymbol{\omega}=\nabla\times\mathbf{u}$ is then

$$\omega=-\tilde{\omega}(1,1,1)+O(|\mathbf{x}|^2) , \tag{41}$$

where $\tilde{\omega}=b-c$. Also, the dissipation rate per unit mass is given by

$$\varepsilon\equiv\frac{1}{2}\nu\left[\frac{\partial u_i}{\partial x_j}+\frac{\partial u_j}{\partial x_i}\right]^2=3\nu\sigma^2+O(|\mathbf{x}|^2) , \tag{42}$$

where $\sigma=b+c$. In the cylindrical coordinates introduced in Sec. III, Eqs. (40) and (41) can be written, respectively, as

$$\mathbf{u}=\begin{bmatrix}u_r\\ u_\theta\\ u_z\end{bmatrix}=\begin{bmatrix}-\sigma r/2\\ -(\sqrt{3}/2)\tilde{\omega}r\\ \sigma z\end{bmatrix}+O(|\mathbf{x}|^3) \tag{43}$$

and

$$\omega=-\sqrt{3}\tilde{\omega}\hat{\mathbf{z}}+O(|\mathbf{x}|^2) . \tag{44}$$

Because of the symmetry of the flow, the pressure term can be expanded in the form

$$p=-d(x^2+y^2+z^2)-e(xy+yz+zx)+O(|\mathbf{x}|^4) . \tag{45}$$

Equation (45) implies that the pressure Hessian π_{ij} has diagonal as well as off-diagonal elements, all of which, in general, can be important in supporting a finite-time singularity. Whereas the diagonal elements, each proportional to d, are determined self-consistently from local properties of the velocity [i.e., they directly enter the differential form of Poisson's equation $\nabla^2 p = -\nabla\cdot(\mathbf{u}\cdot\nabla\mathbf{u})$], the off-diagonal elements, each proportional to e, are determined by the global properties of the flow. If we transform to cylindrical coordinates, the pressure tensor becomes diagonal but not isotropic. Substituting Eqs. (43) and (45) into Eq. (39a), we obtain the exact equations

$$d = bc \ , \tag{46a}$$

$$\dot{b} + c^2 = e \ , \tag{46b}$$

$$\dot{c} + b^2 = e \ , \tag{46c}$$

where an overdot indicates the derivative with respect to time. For self-consistency, e must be calculated using Eqs. (39b) and (45):

$$e = -\left.\frac{\partial^2 p}{\partial x \partial y}\right|_{x=0} = -\int \frac{\partial^2 G(\mathbf{x},t)}{\partial x \partial y}\frac{d^3x}{|\mathbf{x}|} \ , \tag{47}$$

where $G = \nabla\cdot(\mathbf{u}\cdot\nabla\mathbf{u})/4\pi$. Equations (46b) and (46c) can be rewritten as

$$\dot{\tilde{\omega}} = \bar{\omega}\sigma \ , \tag{48a}$$

$$\dot{\sigma} = 2e - (\tilde{\omega}^2 + \sigma^2)/2 \ . \tag{48b}$$

From the form of Eqs. (48a) and (48b), it is easy to see that finite-time singularities in $\tilde{\omega}$ and σ can occur due to the presence of the nonlinear terms for many different functional forms of e. We discuss some interesting examples below.

We consider the case in which the flow is assumed to have the self-similar form

$$\mathbf{u}(\mathbf{x},t) = b[(y,z,x) + a\mathbf{v}_1(\mathbf{x}/a)] + c[(z,x,y) + a\mathbf{v}_2(\mathbf{x}/a)] \ , \tag{49}$$

where a is a function of time only and $\mathbf{v}_1$ and $\mathbf{v}_2$ are two dimensionless functions with no first-order terms in their Taylor expansions. Then it can be shown by dimensional analysis of Eq. (45) that

$$e = c_1 b^2 + c_2 bc + c_3 c^2 \ , \tag{50}$$

where c_1, c_2, c_3 are constants that can be determined once the functional forms of $\mathbf{v}_1$ and $\mathbf{v}_2$ are known. There are numerous choices one can make for c_i and the initial values of b and c that yield singular solutions for Eqs. (46b) and (46c). For example, if we take

$$c_1 = 1, \quad b(0) = 1/t_c, \quad c(0) = 0 \ , \tag{51}$$

then we get

$$b = \tilde{\omega} = \sigma = 1/(t_c - t) \ , \tag{52}$$

which is the solution obtained in Sec. III.

We now consider slightly more general solutions. Let us assume that the flow is quasi-self-similar, i.e., $\mathbf{u}$ can be approximated by Eq. (49) plus a slowly evolving background flow so that

$$e = c_1 b^2 + c_2 bc + c_3 c^2 + c_4 b + c_5 c + c_6 \ , \tag{53}$$

where c_i are slowly varying functions of time (much before the blowup time t_c). As the solutions evolve from different smooth initial conditions to a self-similar asymptotic state, we find that there are two types of singular solutions:

(i) $\sigma = \dfrac{1}{t_c - t}, \quad \tilde{\omega} \neq 0$

(ii) $\sigma = \dfrac{k}{t_c - t}, \quad |\tilde{\omega}/\sigma| \to 0 \ .$

This can be seen by substituting $b = k_1/(t_c - t)$ and $c = k_2/(t_c - t)$ into Eq. (50) [or (53)], whereupon using Eq. (46b) [or (46c)], we get

$$(k_1 - k_2)(k_1 + k_2 - 1) = 0 \ . \tag{54}$$

Type (i) and (ii) singularities are realized, respectively, when the second and first factors are set equal to zero.

Type (i) and (ii) singularities can also be obtained by considering a self-similar solution to the Euler equation of the form

$$\mathbf{u}(\mathbf{x},t) = \frac{a(t)}{t_c - t}\mathbf{V}(\mathbf{x}/a), \quad a(t) = \mu(t_c - t)^p \ . \tag{55}$$

The curl of Eq. (39a) yields

$$\nabla\times[(1-p)\mathbf{V}(\mathbf{X}) + (p\mathbf{X} + \mathbf{V})\cdot\nabla\mathbf{V}] = 0 \ , \tag{56}$$

where $\mathbf{X} = \mathbf{x}/a$. Let

$$\mathbf{V} = [BY + CZ, BZ + CX, BX + CY] + O(|\mathbf{X}|^2) \ ; \tag{57}$$

we then have

$$(B - C)(B + C - 1) = 0 \ , \tag{58}$$

which is similar to Eq. (54).

In writing Eq. (55), we have assumed the global existence of a divergence-free function $\mathbf{V}(\mathbf{X})$. One way to test this assumption is by rewriting Eq. (55) in the form of an integral equation

$$\mathbf{V}(\mathbf{X}) = -\frac{1}{4\pi(1-p)}\int \frac{\nabla'\times[\{p\mathbf{X}' + \mathbf{V}(\mathbf{X}')\}\cdot\nabla'\mathbf{V}(\mathbf{X}')]\times(\mathbf{X} - \mathbf{X}')}{|\mathbf{X} - \mathbf{X}'|^3} d\mathbf{X} \ , \tag{59}$$

which may be tractable numerically. This is a well-posed problem, but we will not try to solve it here.

The model analyzed in Sec. III is a special case of (i) with $c = 0$ and leads to a finite-time vortex singularity at the origin. At first glance, case (ii) may appear to violate the Beale-Kato-Majda constraint [44] that the maximum vorticity of a flow with an algebraic finite-time singularity has to blow up at least at fast as $(t_c - t)^{-1}$. However, this

is not necessarily so. Since both b and c blow up at a rate proportional to $(t_c-t)^{-1}$, it is clear by inspection of (49) that the maximum vorticity due to higher-order terms must also blow up as $(t_c-t)^{-1}$ if $\nabla\times(\mathbf{v}_1+\mathbf{v}_2)\neq 0$, which is generally the case. (A more precise statement is given at the end of Sec. VII.)

Whereas the model considered in Ref. [28] (and Sec. III) is a singular solution of type (i), the singularities obtained in the high-symmetry numerical experiments of Refs. [45] and [46] are examples of type (ii). Consider, for example, the singularity obtained by Kerr in a numerical study of two interacting, antiparallel vortex tubes [45]. Because of the symmetries of Kerr's initial condition, the velocity and vorticity remain zero at the point $\mathbf{x}_c$ where the singularity eventually develops as the two tubes approach each other. The Taylor expansion of the flow about $\mathbf{x}_c$ can be written as

$$\mathbf{v}=[a_1x,a_2y,-(a_1+a_2)z]+O(|\mathbf{x}|^2)\ , \tag{60}$$

where a_1,a_2,a_3 are constants. We note that the $|\mathbf{x}|^2$ terms in Eq. (60) do not affect the analysis given above as long as $\mathbf{v}(\mathbf{x}=0)$ is fixed at zero. Then it follows that in Kerr's geometry a self-similar singular solution should occur with maximum vorticity blowing up as $(t_c-t)^{-1}$.

We conclude with some numerical examples. With e given by (50), we plot in Fig. 1 an example of a type (i) singularity, with $c_1=0.9$, $c_2=-0.9$, $c_3=-0.99$, and the initial conditions $b(0)=1$ and $c(0)=0.5$. In Fig. 2 we plot an example of a type (ii) singularity, with $c_1=10.1$, $c_2=-1.3$, $c_3=-1.45$, and the initial conditions $b(0)=1$ and $c(0)=1.0e-06$.

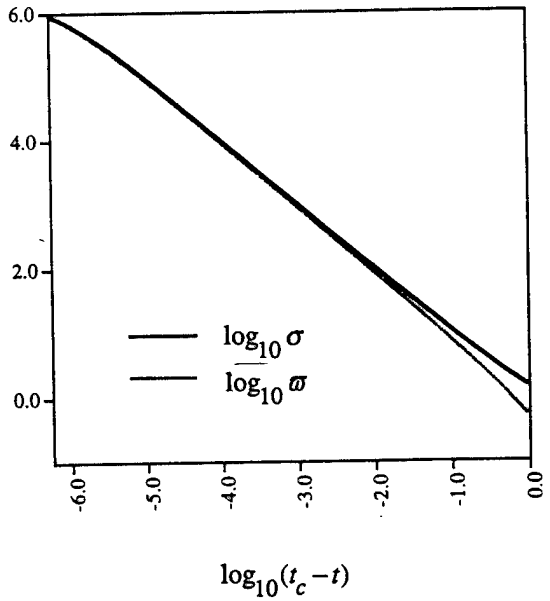

FIG. 1. Example of a singular solution of type (i) where both ϖ and σ are proportional to $1/(t_c-t)$ using constants $c_1=0.9$, $c_2=-0.9$, and $c_3=-0.99$ with initial values $b(0)=1$ and $c(0)=0.5$. The constant t_c is found to be 0.9495.

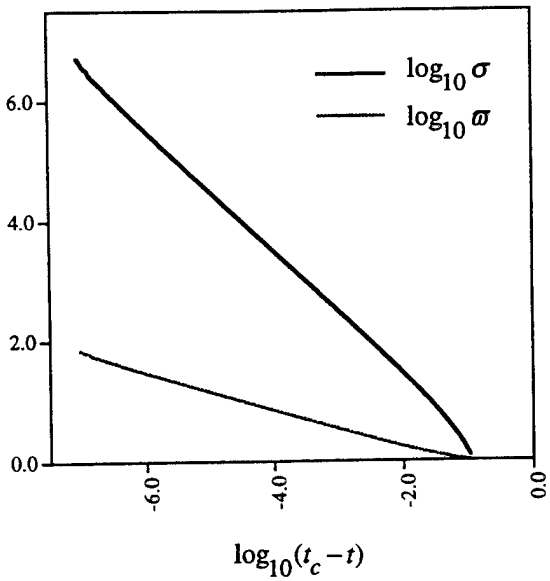

FIG. 2. Example of a singular solution of type (ii) where $\sigma=1/(t_c-t)$ and $\tilde{\omega}/\sigma\to 0$ using constants $c_1=10.1$, $c_2=-1.3$, and $c_3=-1.45$ with initial values $b(0)=1$ and $c(0)=10^{-6}$. The constant t_c is found to be 0.1153.

V. SINGULARITY DYNAMICS IN THE COMPLEX SPATIAL DOMAIN

As mentioned in Sec. I, Tanveer and Speziale [29] have proposed a method for investigating the singularity dynamics of the Euler equation in the complex physical domain. In this method, a class of smooth and bounded initial conditions in the real physical domain is analytically continued into the complex unphysical domain where there are complex singularities. Then the question of finite-time singularities in the real physical domain can be reduced to the question of whether the singularities in the complex domain reach the real domain in finite time.

The method proposed by Tanveer and Speziale [29] is based on some important assumptions that have not been shown to be valid in general. The method has been used successfully to obtain sufficient conditions for stability for a class of three-dimensional Euler flows identical to those found earlier by WKB methods [24,25]. In what follows, we suggest that some of the assumptions in Ref. [29] can be relaxed and that the method can be applied to our symmetric initial condition to obtain useful insights. There are open mathematical questions that we do not settle here, but since we state our assumptions clearly, it is our hope that the validity of the framework of Ref. [29] in the context of our model can be determined by further work.

For the purpose of this section, we consider $\mathbf{x}$ to be a complex variable and consider initial conditions for which the velocity and vorticity are real and analytic everywhere for real $\mathbf{x}$. However, the initial condition has complex singularities at the surface

$$d(\mathbf{x},0)=0\ , \tag{61}$$

where $d(\mathbf{x},0)$ is a real and positive analytic function for real $\mathbf{x}$. Following Ref. [29], we take the initial conditions

for the velocity and the vorticity to be of the form

$$\mathbf{u}(\mathbf{x},0)=\mathbf{v}_s(\mathbf{x},0)+f[d(\mathbf{x},0)]\mathbf{q}(\mathbf{x},0) \tag{62}$$

and

$$\boldsymbol{\omega}(\mathbf{x},0)=\boldsymbol{\omega}_s(\mathbf{x},0)+f'[d(\mathbf{x},0)]\mathbf{p}(\mathbf{x},0)\ , \tag{63}$$

respectively, where $f(d)$ is a function to be specified. Tanveer and Speziale assume that

$$f(d)=d^{\alpha},\quad 0<\alpha<1\ . \tag{64}$$

In (62) and (63), $\mathbf{v}_s(\mathbf{x},0),\boldsymbol{\omega}_s(\mathbf{x},0),\mathbf{q}(\mathbf{x},0)$, and $\mathbf{p}(\mathbf{x},0)$ are real and analytic functions of $\mathbf{x}$ that obey the equations $\nabla\cdot\mathbf{v}_s(\mathbf{x},0)=0,\nabla\times\mathbf{v}_s(\mathbf{x},0)=\boldsymbol{\omega}_s(\mathbf{x},0)$, and $\nabla\times[f\mathbf{q}(\mathbf{x},0)]=f'\mathbf{p}(\mathbf{x},0)$. It is claimed in Ref. [29] that for $t>0$, the complex solutions

$$\mathbf{u}(\mathbf{x},t)=\mathbf{v}_s(\mathbf{x},t)+f[d(\mathbf{x},t)]\mathbf{q}(\mathbf{x},t)\ , \tag{65}$$

$$\boldsymbol{\omega}(\mathbf{x},t)=\boldsymbol{\omega}_s(\mathbf{x},t)+f'[d(\mathbf{x},t)]\mathbf{p}(\mathbf{x},t) \tag{66}$$

may be constructed with

$$\nabla\cdot\mathbf{v}_s(\mathbf{x},t)=0\ , \tag{67}$$

$$\nabla\times\mathbf{v}_s(\mathbf{x},t)=\boldsymbol{\omega}_s(\mathbf{x},t)\ , \tag{68}$$

$$\frac{\partial\boldsymbol{\omega}_s}{\partial t}+\mathbf{v}_s\cdot\nabla\boldsymbol{\omega}_s=\boldsymbol{\omega}_s\cdot\nabla\mathbf{v}_s\ , \tag{69}$$

with $\mathbf{p}$ and $\mathbf{q}$ satisfying Eqs. (21), (22), and (25) of Ref. [29]. The singular surface $d(\mathbf{x},t)=0$ is evolved according to the equation

$$\frac{\partial d(\mathbf{x},t)}{\partial t}+\mathbf{v}_s\cdot\nabla d(\mathbf{x},t)=0\ . \tag{70}$$

Equation (70) implies that the singularities of $\boldsymbol{\omega}(\mathbf{x},t)$ for $t>0$ are determined by the relation $d(\mathbf{x},t)=0$. Thus Tanveer and Speziale obtain the surprising result that the location of the singularities in the complex domain can be determined from the smooth velocity field $\mathbf{v}_s(\mathbf{x},t)$ without the need to solve for $\mathbf{p}(\mathbf{x},t)$ and $\mathbf{q}(\mathbf{x},t)$, assuming that they exist. Tanveer and Speziale demonstrate that no complex singularity can reach the real domain in finite time if the variables $\mathbf{p},\mathbf{q},\mathbf{v}_s$, or $\boldsymbol{\omega}_s$ are smooth and develop no spontaneous singularities.

There has been some criticism [47] of Eq. (70) on the ground that it does not capture the nonlocal effect of pressure. We point out that whether that is so depends on the choice of $\mathbf{v}_s(\mathbf{x},0)$ and $\boldsymbol{\omega}_s(\mathbf{x},0)$, which influence the solutions of Eqs. (67) and (68) and hence the solution for $d(\mathbf{x},t)$. Furthermore, the nonlocal effect of pressure also enters into the equations for $\mathbf{p}$ and $\mathbf{q}$ that are assumed to exist.

A key difference between the considerations of Ref. [29] and what follows here is that we choose an exact but singular solution of the Euler equation for $\mathbf{v}_s$ and $\boldsymbol{\omega}_s$, i.e.,

$$\mathbf{v}_s=\frac{1}{t_c-t}\begin{bmatrix}-r/2\\ -\sqrt{3/2}r\\ z\end{bmatrix}\ , \tag{71}$$

$$\boldsymbol{\omega}_s=-\frac{\sqrt{3}}{t_c-t}\begin{bmatrix}0\\0\\1\end{bmatrix}\ . \tag{72}$$

Though motivated by our real analysis in Sec. III, the choice (71) does not rely in any way on the validity of that analysis. In order that $\mathbf{u}$ and $\boldsymbol{\omega}$ remain bounded initially for large $|\mathbf{x}|$, we assume that $f(d(\mathbf{x})\to\infty)\to 1$ and furthermore that

$$\mathbf{q}(\mathbf{x},0)=\mathbf{q}_1(\mathbf{x},0)-\mathbf{v}_s(\mathbf{x},0)\ , \tag{73}$$

where $\mathbf{q}_1(\mathbf{x},0)$ is smooth and decays to zero as $|\mathbf{x}|\to\infty$. We choose

$$f(d)=1-\exp(-d^{\alpha}),\quad 0<\alpha<1\ . \tag{74}$$

Equation (74) gives $f(0)=0$, $f'(0)\to\infty$, which enables us to carry through with the approach of Ref. [29], assuming that there exist well-behaved solutions $\mathbf{p}$ and $\mathbf{q}$, satisfying Eqs. (21), (22), and (25) of Ref. [29], such that the complex singularities in (66) will not be canceled. We are not able to show definitely whether this assumption is true, because it involves proofs of existence of solutions to the cited partial differential equations of Ref. [29]. Further analytical and numerical work will be required to settle this issue.

The relation $d(\mathbf{x},t)=0$ is satisfied by complex points $\mathbf{x}(t)$, which obey the characteristic equation

$$\frac{d\mathbf{x}(t)}{dt}=\mathbf{v}_s(\mathbf{x}(t),t)\ , \tag{75}$$

subject to the initial condition (61). A solution of this equation is

$$d(\mathbf{x},t)=\frac{t_c}{t_c-t}\left[r^2+z^2\left[\frac{t_c-t}{t_c}\right]^3+b^2(0)\left[\frac{t_c-t}{t_c}\right]\right] \tag{76}$$

for $t<t_c$. We see that at $t=0$, complex singularities in vorticity occur on a cylindrical manifold, defined by

$$r^2+z^2+b^2(0)=0\ . \tag{77}$$

For $t>0$, the coordinates of the movable singularities are determined by setting the right-hand side of Eq. (76) to zero. On the $z=0$ plane, we have

$$r(t)=ib(t)=ib(0)[(t_c-t)/t_c]^{1/2}\ . \tag{78}$$

Equation (78) suggests, remarkably, that not only do the complex singularities on the $z=0$ plane reach the origin (stagnation point) in finite time, but that they do so at a rate proportional to $(t_c-t)^{1/2}$. [From Eqs. (70) and (76) it might appear, upon first glance, that $d(\mathbf{x}=0,t)=$const for all time. However, this would not be a correct inference from Eq. (76), which holds only for $t<t_c$. At $t=t_c$, the velocity $\mathbf{v}_s(\mathbf{x}=0,t)$ is no longer zero but becomes singular.] Here $b(t)$ defines a small space scale in the complex spatial domain that is distinct from the collapsing small scale of the flow $\mathbf{v}_s$. These results are consistent with the results obtained in Sec. III.

VI. REMARKS ON THE CONSTANTIN-FEFFERMAN THEOREM

Constantin [30,31] and Fefferman [31] have recently proved that "if the direction of vorticity is sufficiently well behaved in regions of high vorticity magnitude, then the solution is smooth." (The words "sufficiently well behaved" have a technical meaning that we consider carefully later.) This result imposes strong constraints on the possible singularities of the Euler equation. The crux of the theorem is in the relation [30]

$$\frac{d|\omega|}{dt}=\alpha_C|\omega| \ , \tag{79}$$

where

$$\alpha_C(\mathbf{x},t)=\frac{3}{4\pi}\mathrm{P}\int\frac{d\mathbf{y}}{|\mathbf{y}|^3}D(\hat{\mathbf{y}},\hat{\boldsymbol{\omega}}(\mathbf{x}+\mathbf{y},t),\hat{\boldsymbol{\omega}}(\mathbf{y},t)) \ , \tag{80}$$

$\hat{\boldsymbol{\omega}}\equiv\boldsymbol{\omega}/|\boldsymbol{\omega}|$ is the unit vorticity vector and D is

$$D(\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3)=(\mathbf{e}_1\cdot\mathbf{e}_3)[\mathrm{Det}(\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3)] \ . \tag{81}$$

Since D vanishes if any of the column vectors in the determinant are parallel or antiparallel, spatial alignment or antialignment depletes the growth of vorticity, eliminating the possibility of a singularity. This is what happens in two dimensions globally. In three dimensions, in the neighborhood of a potential singularity, if D is too small, then it is not possible to support the singularity. This underscores the importance of infinite spatial gradients or small scales.

In view of the relation (79), the qualitative conclusions of the Constantin-Fefferman theorem are physically compelling. The question then is how our model measures up against this theorem. In Secs. III and V, we have established the growth of small scales. In particular, Contantin [30] shows that if the vorticity blows up self-similarly as

$$\omega(\mathbf{x},t)=\frac{1}{t_c-t}\mathbf{A}\left[\frac{\mathbf{x}}{a}\right]\equiv\frac{1}{t_c-t}\mathbf{A}(\mathbf{X}) \ , \tag{82}$$

where $a=\mu(t_c-t)^p$, then $p\geq 2/5$. This condition is satisfied by our locally self-similar solution. Furthermore, since the velocity blows up in the shrinking middle region, it is clear that our model also does not contradict a variant of the Constantin-Fefferman theorem, which states that there can be no singularity if the velocity does not blow up and the vorticity is locally absolutely integrable [31].

While accord with the theorem is reassuring, we discuss a technical issue that has to do with the words "sufficiently well behaved" in the first line of this section. In order to control the size of a "dangerous term" [31], the theorem needs the following assumption: There exist two positive constants Ω_C and ρ_C such that for every pair of two locations $|\mathbf{x}_1|$ and $|\mathbf{x}_2|$, whenever $|\boldsymbol{\omega}(\mathbf{x}_1,t)|>\Omega_C$ and $|\boldsymbol{\omega}(\mathbf{x}_2,t)|>\Omega_C$, we have

$$|\sin\phi(\mathbf{x}_1,\mathbf{x}_2,t)|\leq|\mathbf{x}_1-\mathbf{x}_2|/\rho_C \ , \tag{83}$$

where $\phi(\mathbf{x}_1,\mathbf{x}_2,t)$ is the angle between $\boldsymbol{\omega}(\mathbf{x}_1,t)$ and $\boldsymbol{\omega}(\mathbf{x}_2,t)$.

We now show that a self-similar vorticity field of the form (82) does not satisfy assumption (83). To see this, we choose two dimensionless vectors $\mathbf{X}_1$ and $\mathbf{X}_2$ such that $\mathbf{X}_1\neq\mathbf{X}_2$, $\mathbf{A}(\mathbf{X}_1)\|\mathbf{A}(\mathbf{X}_2)|\neq 0$, and $|\sin\phi|\neq 0$. We consider two corresponding locations in real space, $\mathbf{x}_1=a_m\mathbf{X}_1$ and $\mathbf{x}_2=a_m\mathbf{X}_2$, where $a_m=\mu(t_c-t_m)^p$. Then at $t=t_m$, we have

$$|\boldsymbol{\omega}(\mathbf{x}_i,t_m)|=\left[\frac{\mu}{a_m}\right]^{1/p}|\mathbf{A}(\mathbf{X}_i)| \ , \tag{84}$$

where $i=1,2$ and

$$\frac{|\sin\varphi(\mathbf{x}_1,\mathbf{x}_2,t_m)|}{|\mathbf{x}_1-\mathbf{x}_2|}=\frac{1}{a_m|\mathbf{X}_1-\mathbf{X}_2|}\left[1-\frac{|\mathbf{A}(\mathbf{X}_1)\cdot\mathbf{A}(\mathbf{X}_2)|^2}{|\mathbf{A}(\mathbf{X}_1)|^2|\mathbf{A}(\mathbf{X}_2)|^2}\right]^{1/2} \ . \tag{85}$$

From (85), we see that for every choice of Ω_C and ρ_C, we can find a small enough a_m such that $|\boldsymbol{\omega}(\mathbf{x}_1,t)|>\Omega_C$ and $|\boldsymbol{\omega}(\mathbf{x}_2,t)|>\Omega_C$, but $|\sin\phi(\mathbf{x}_1,\mathbf{x}_2,t_m)|>|\mathbf{x}_1-\mathbf{x}_2|/\rho_C$. We note that this holds even for the case $\mathbf{A}(\mathbf{X}=0)=0$, but $\mathbf{A}(\mathbf{X}\neq 0)\neq 0$. Thus assumption (83) is violated. Since the assumption of local self-similarity is inherent in our treatment, it is questionable whether the Constantin-Fefferman theorem is relevant to our model.

Before we conclude this section, we comment on case (ii), discussed in Sec. III. The argument given in the preceding paragraph enables us to see why the Beale-Kato-Majda theorem [44] is not violated in those cases where the velocity null coincides with a vorticity null, examples of which are discussed in Sec. III. From (82), it follows that $\max|\boldsymbol{\omega}(\mathbf{x},t)|=(t_c-t)^{-1}|\mathbf{A}(\mathbf{X}_m)|$ exists even when $\mathbf{A}(\mathbf{X}=0)=0$, but $\mathbf{A}(\mathbf{X}\neq 0)\neq 0$. Since $\mathbf{X}_m$ is independent of t, it follows that $\max|\boldsymbol{\omega}(\mathbf{x},t)|$ blows up as $(t_c-t)^{-1}$, consistent with the Beale-Kato-Majda constraint.

VII. EFFECT OF VISCOSITY

A. Connection with Lundgren's model

In Lundgren's model, a two-dimensional flow with vorticity $\omega_2(r,\theta,t)$ is placed in an axisymmetry straining flow with velocity components $u_z=s(t)z$ and $u_r=-s(t)r/2$, where $s(t)$ is the strain rate. This is shown to produce a three-dimensional axially strained flow with the vorticity

$$\omega_z(r,\theta,t)=S(t)\omega_2[\xi,\theta,T] \ , \tag{86}$$

where the variables S, ξ, and T are defined, respectively, by Eqs. (30), (31a), and (31d). At $t=0$, the vorticity of the three-dimensional flow is equal to the vorticity of the two-dimensional flow.

A connection can be established between the present model and Lundgren's by using Eqs. (27). If initially,

$$\frac{\partial}{\partial r_0}\sim\frac{1}{r_0}\frac{\partial}{\partial\theta_0}\sim\frac{\partial}{\partial z_0} \ , \tag{87}$$

then as $t\rightarrow t_c$, we have

$$\frac{\partial}{\partial r}=\frac{\partial r_0}{\partial r}\frac{\partial}{\partial r_0}=\left[\frac{t_c}{t_c-t}\right]^{1/2}\frac{\partial}{\partial r_0}$$
$$\sim\frac{1}{r}\frac{\partial}{\partial\theta}=\left[\frac{t_c}{t_c-t}\right]^{1/2}\frac{1}{r_0}\frac{\partial}{\partial\theta_0}$$
$$\gg\frac{\partial}{\partial z}=\left[\frac{t_c-t}{t_c}\right]\frac{\partial}{\partial z_0}\ ,$$

which justifies setting $\partial/\partial z=0$.

We seek solutions of the Navier-Stokes equation

$$\frac{\partial\boldsymbol{\omega}}{\partial t}+\mathbf{u}\cdot\nabla\boldsymbol{\omega}=\boldsymbol{\omega}\cdot\nabla\mathbf{u}+\nu\nabla^2\boldsymbol{\omega}\ , \tag{88}$$

where ν is the viscosity. Near the origin where the vortex singularity occurs for inviscid flows, the vorticity has a dominant z component that is independent of z. We write

$$\omega_z=-\sqrt{3}s(t)+\alpha(r,\theta,t)\ . \tag{89}$$

The self-consistent velocity associated with this vorticity can be approximated as

$$u_z=s(t)z+\beta_z(r,\theta,t)\ , \tag{90}$$
$$u_r=-s(t)r/2+\beta_r(r,\theta,t)\ , \tag{91}$$
$$u_\theta=-(\sqrt{3}/2)s(t)r+\beta_\theta(r,\theta,t)\ . \tag{92}$$

Since $\nabla\cdot\mathbf{u}=0$, we get

$$\frac{\partial}{\partial r}(r\beta_r)+\frac{\partial\beta_\theta}{\partial\theta}=0\ , \tag{93}$$

which implies that β_r and β_θ are derivable from a stream function $\psi(r,\theta,t)$, i.e.,

$$\beta_r=\frac{1}{r}\frac{\partial\psi}{\partial\theta},\quad \beta_\theta=-\frac{\partial\psi}{\partial r}\ . \tag{94}$$

From the z component of the equation $\boldsymbol{\omega}=\nabla\times\mathbf{u}$, we then get

$$\alpha=-\nabla^2\psi\ . \tag{95}$$

The dynamical equation for α follows from the z component of the Navier-Stokes equation (88)

$$\frac{\partial\alpha}{\partial t}+\left[\beta_r-\frac{sr}{2}\right]\frac{\partial\alpha}{\partial r}+\left[\frac{\beta_\theta}{r}-\frac{\sqrt{3}s}{2}\right]\frac{\partial\alpha}{\partial\theta}$$
$$=\sqrt{3}\frac{\partial s}{\partial t}+s(-\sqrt{3}s+\alpha)+\nu\nabla^2\alpha\ . \tag{96}$$

As shown by Lundgren [32], by transforming to the variables defined by Eqs. (29)–(31), it is possible to construct solutions to Eq. (96) from solutions of the two-dimensional equation

$$\frac{\partial\alpha_2}{\partial T}+\frac{1}{\xi}\left[\frac{\partial\psi_2}{\partial\vartheta}\frac{\partial\alpha_2}{\partial\xi}-\frac{\partial\psi_2}{\partial\xi}\frac{\partial\alpha_2}{\partial\vartheta}\right]-\nu\nabla_\xi^2\alpha_2$$
$$=-\frac{\sqrt{3}}{S^2}\left[s^2-\frac{\partial s}{\partial t}\right]\ , \tag{97}$$

where

$$\psi(r,\theta,t)=\psi_2(\xi,\vartheta,T)\ , \tag{98a}$$
$$\alpha(r,\theta,t)=S(t)\alpha_2(\xi,\vartheta,T)\ , \tag{98b}$$
$$\alpha_2(\xi,\vartheta,T)=-\nabla_\xi^2\psi_2\ , \tag{98c}$$

and

$$\nabla_\xi^2=\frac{\partial^2}{\partial\xi^2}+\frac{1}{\xi}\frac{\partial}{\partial\xi}+\frac{1}{\xi^2}\frac{\partial^2}{\partial\vartheta^2}\ . \tag{99}$$

For inviscid flows, we have shown earlier that $s(t)$ obeys Eq. (23), which yields a finite-time singularity of the vorticity. Then the stretch ratio $S(t)$ becomes infinitely large in finite time. (In contrast, the stretch ratio in Lundgren's model grows exponentially with time.) Due to this tendency for explosive growth, the right-hand side of Eq. (97) eventually becomes much smaller than the left-hand side and the equation reduces to

$$\frac{\partial\alpha_2}{\partial T}+\frac{1}{\xi}\left[\frac{\partial\psi_2}{\partial\vartheta}\frac{\partial\alpha_2}{\partial\xi}-\frac{\partial\psi_2}{\partial\xi}\frac{\partial\alpha_2}{\partial\vartheta}\right]-\nu\nabla_\xi^2\alpha_2=0\ . \tag{100}$$

The presence of even a small but finite viscosity thwarts the formation of the vortex singularity. This can be seen by introducing a viscous correction to Eq. (97). Using Eq. (21) to calculate the term $\nu\nabla^2\omega_z\cong 4\sqrt{3}s\nu/a^2$, we can rewrite Eq. (23) as

$$\frac{\partial s}{\partial t}-s^2+\frac{4s\nu}{a^2}\cong 0\ . \tag{101}$$

Thus, in the presence of viscosity, the strain saturates at a value s_c, which can be obtained by setting $\partial s/\partial t=0$. We then get the dissipation scale

$$a_d\cong 2\left[\frac{\nu}{s_c}\right]^{1/2}\ . \tag{102}$$

The dissipation rate per unit mass (for the local flow) is

$$\varepsilon\equiv\frac{1}{2}\nu\left[\frac{\partial u_i}{\partial x_j}+\frac{\partial u_j}{\partial x_i}\right]^2\cong 3\nu s_c^2\ . \tag{103}$$

Combining Eqs. (102) and (103), we get

$$a_d\cong 2(3\nu^3/\varepsilon)^{1/4}\approx\eta\equiv(\nu^3/\varepsilon)^{1/4}\ , \tag{104}$$

where η is the Kolmogorov scale.

B. The spiral structure

It is clear by inspection of the initial conditions (Sec. II) that a spiral structure occurs naturally in the strained vortex solution. This spiral structure is also manifest in Eq. (28) as we go beyond the leading-order axisymmetric solution. We now follow Lundgren's method and seek two-dimensional vortex spiral solutions $\alpha_2(r,\theta,t)$ of the equation

$$\frac{\partial \alpha_2}{\partial t}+\frac{1}{r}\left[\frac{\partial \psi_2}{\partial \theta}\frac{\partial \alpha_2}{\partial r}-\frac{\partial \psi_2}{\partial r}\frac{\partial \alpha_2}{\partial \theta}\right] =\nu\left[\frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}+\frac{1}{r^2}\frac{\partial^2}{\partial \theta^2}\right]\alpha_2 . \quad (105)$$

Once the two-dimensional solution (i.e., the solution for which the axial strain $s=0$) to Eq. (105) is known, we can write down the three-dimensional axially strained solution by using Eq. (98b). For the two-dimensional equation (105), we seek solutions of the form

$$\alpha_2(r,\theta,t)=A_0(r,t)+\sum_{\pm}A_{\pm}(r,t)\exp\{\pm i3(\theta-\alpha t)\} , \quad (106)$$

where $\tilde{\alpha}=\beta_\theta/r$ is a slowly varying function of time. We write the stream function $\psi_2(r,\theta,t)$ as

$$\psi_2(r,\theta,t)=\Psi_0(r)+\Psi_1(r,\theta,t) , \quad (107)$$

where $\tilde{\alpha}\cong-(1/r^2)(d\Psi_0/dr)$. Hence Eq. (105) gives

$$\frac{\partial \alpha_2}{\partial t}+\tilde{\alpha}\frac{\partial \alpha_2}{\partial \theta}=\nu\left[\frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}+\frac{1}{r^2}\frac{\partial^2}{\partial \theta^2}\right]\alpha_2 , \quad (108)$$

where

$$A_0=\frac{1}{r}\frac{d}{dr}(r^2\tilde{\alpha}) . \quad (109)$$

The zeroth harmonic of the vorticity satisfies the heat equation

$$\frac{\partial A_0}{\partial t}=\nu\left[\frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}\right]A_0 , \quad (110)$$

which has the solution $A_0=(1/t)\exp(-r^2/4\nu t)$. Hence A_0, as well as $\tilde{\alpha}$, has a slow decay time τ, defined by $\langle\tilde{\alpha}\rangle\tau=\langle\tilde{\alpha}\rangle a^2/\nu$, where $\langle\tilde{\alpha}\rangle$ is a characteristic magnitude of $\tilde{\alpha}$. As $\nu\to0$, the harmonics $A_\pm$ obey the equation [32]

$$\frac{\partial A_\pm}{\partial t}\cong-9\nu\tilde{\alpha}'^2t^2A_\pm \quad (111)$$

and have the asymptotic behavior

$$A_\pm\cong f_\pm(r)\exp(-3\nu\tilde{\alpha}'^2t^3) , \quad (112)$$

where a prime means the derivative with respect to the argument and $f_\pm(r)$ represent arbitrary functions of r. Using Eq. (112), we can define a decay time for the harmonics by the relation

$$\langle\tilde{\alpha}\rangle\tau_\pm=(\langle\tilde{\alpha}\rangle a^2/\nu)^{1/3} . \quad (113)$$

Since the Reynolds number $\langle\tilde{\alpha}\rangle a^2/\nu$ is large, we have $\tau_\pm\ll\tau_0$, which means that the higher harmonics decay much faster than the zeroth harmonic. Due to nonlinear mode coupling, we expect that other harmonics will be generated though they are not present in the initial conditions. In what follows, we shall neglect the contribution of these higher harmonics.

Using the two-dimensional spiral solution in Eq. (98b), we obtain the axially strained spiral solution, given by

$$\alpha(r,\theta,t)=W_0(r,t)+\sum_{\pm}W_\pm(r,t)\exp\{\pm i3(\tilde{\theta}-\tilde{\alpha}T)\} , \quad (114)$$

where

$$W_0=SA_0(S^{1/2}r)=(S/T)\exp(-Sr^2/4\nu T) \quad (115)$$

and

$$W_\pm=Sf_\pm(S^{1/2}r)\exp[-3\{\tilde{\alpha}'(S^{1/2}r)\}^2\nu T^3] . \quad (116)$$

In Eqs. (115) and (116), the variables S and T will hereafter be evaluated using the strain rate $s=s_c$. Using Eq. (116), the fast decay time $T_\pm$ is seen to be

$$T_\pm=(3\nu\tilde{\alpha}'^2)^{-1/3} . \quad (117)$$

The rapidly oscillating factor $\exp(\mp i3\tilde{\alpha}T)\equiv\exp[\mp iR(r,t)]$ can be used to define a local wave number

$$q\equiv\partial R/\partial r=3S^{1/2}\tilde{\alpha}'T . \quad (118)$$

Using Eq. (117), q can be estimated to have the characteristic value

$$q_c\sim(s_c/\nu)^{1/2}\sim\eta^{-1} , \quad (119)$$

where η is defined by Eq. (101). Thus the Kolmogorov scale is an intrinsic feature of the spiral solution (including the zeroth and higher harmonics).

As discussed by Lundgren, Eq. (118) indicates that since S, as well as T, increases with time, the wave number q also increases with time while $\tilde{\alpha}'$ remains approximately constant. This mechanism of transfer of energy to small scales is associated with the tightening of the spiral structure, brought about by differential rotation and axial straining.

C. The Kolmogorov spectrum

As the vorticity intensifies to extremely large values, viscosity intervenes, thwarts the formation of the (inviscid) finite-time singularities, and causes transition to turbulence. It has been shown by Lundgren [32] that the energy spectrum of this system is given by

$$E(k)=\frac{C}{k^2}\int_0^{T_c}S^{1/2}(T)F_2(K,T)dT , \quad (120)$$

where $K\equiv kS^{-1/2}(T)$, T_c is the lifetime of a vortex segment of initial length l_0, and C is a constant defined as $C=2\pi^2l_0N_c/L^3$, where N_c is the rate of creation of vortices. (The parameter l_0N_c/L^3 represents the rate of creation of vortex length per unit volume.) The function F_2 is the enstrophy spectrum, defined by the equation,

$$F_2(k,t)=k\int_0^{2\pi}|\bar{\alpha}_2(k\cos\theta_k,k\sin\theta_k,t)|^2d\theta_k , \quad (121)$$

where

$$\bar{\alpha}_2(\mathbf{k},t)=(2\pi)^{-2}\int\int dr\,d\theta\,r\alpha_2(r,\theta,t)\exp(-i\mathbf{k}\cdot\mathbf{r}) . \quad (122)$$

Here $\mathbf{k}$ is defined on the two-dimensional (r,θ) plane and θ_k is the angle between the unit vectors $\hat{\mathbf{k}}$ and $\hat{\mathbf{x}}$.

The enstrophy spectrum (121) can be calculated by using the solutions of Eqs. (111) and (112). It has been shown [32–34] that the zeroth harmonic gives $E_0(k)\sim k^{-3}$. Since we anticipate that this will be subdominant to the contribution of the spirals to the energy spectrum for large k, we consider only the contribution of the spirals, described by the asymptotic solutions (112). Following the steps described in Ref. [32], we get

$$F_2(k,t)=\sum_{\pm} k\exp(-6\nu\tilde{\alpha}'^2t^3)\times\left|\int_0^\infty dr\, rf_\pm(r)J_{\pm3}(kr)\exp(\mp3i\tilde{\alpha}t)\right|^2. \tag{123}$$

For large k and t, using the asymptotic expression for the Bessel function, we obtain a rapidly varying function in the integrand of the form $\exp\{-i(kr+3\tilde{\alpha}t)\}$. The integral in Eq. (123) can then be evaluated by the method of stationary phase. The result is

$$\left|\int_0^\infty dr\, rf_\pm(r)J_\pm(kr)\exp(\mp3i\tilde{\alpha}t)\right|^2\cong\frac{r_s|f_\pm(r_s)|^2}{2k\tilde{\alpha}''(r_s)t}, \tag{124}$$

where $r_s=r_s(k/t)$ is given by the stationary-phase condition $k+3\tilde{\alpha}'(r_s)t=0$. It is now apparent, by inspection of Eqs. (123) and (124), that the enstrophy spectrum $F_2(k,t)$ has the self-similar form $t^{-1}G(k/t)$ in the inertial range $L^{-1}\ll k\ll\eta^{-1}$. This similarity form for the enstrophy, which is preserved when dissipation can be neglected, is responsible for the Kolmogorov spectrum [32,34]. Inserting Eqs. (123) and (124) in Eq. (120), we obtain the contribution of the spirals to the energy spectrum

$$E_\pm(k)=C_K\varepsilon^{2/3}k^{-5/3}\exp(-\tfrac{2}{3}\eta^2k^2), \tag{125}$$

with the coefficient

$$C_K=\frac{(\frac{2}{3})^{2/3}}{\Gamma(\frac{2}{3})}\frac{\varepsilon_\pm}{\varepsilon^{2/3}\nu^{1/3}s_c^{2/3}}. \tag{126}$$

Using the definitions

$$\xi_s\equiv S^{1/2}r_s,\quad\left|\frac{\tilde{\alpha}'}{\tilde{\alpha}''}\right|dT\equiv\frac{2}{3}Td\xi_s \tag{127}$$

and writing $s_c\approx S/T$, we obtain

$$\varepsilon_\pm=4\pi C\nu^{1/3}s_c\frac{\Gamma(\frac{2}{3})}{2^{2/3}}\left[\int_0^\infty d\xi_s\frac{\xi_s|f_\pm(\xi_s)|^2}{|\tilde{\alpha}'(\xi_s)|^{4/3}}\right]. \tag{128}$$

At first sight, C_K may appear to be model dependent, since Eq. (128) suggests so. However, since $E_\pm(k)\gg E_0(k)$ and $\varepsilon_\pm\approx\varepsilon$, we note that Eq. (125) reduces to

$$C_K\approx\frac{(\frac{2}{3})^{2/3}}{\Gamma(\frac{2}{3})}\left[\frac{\varepsilon}{\nu s_c^2}\right]^{1/3}. \tag{129}$$

If we recall that $\varepsilon\cong3\nu s_c^2$, we obtain a constant $C_K\approx0.8$.

VIII. SUMMARY AND DISCUSSION

In this paper, we have built upon a recent model of three-dimensional Euler flows that yields a finite-time vortex singularity [28]. In Ref. [28] as well as here, we have emphasized that *geometrical* features of an inviscid flow have a strong role in determining where unbounded local vorticity growth occurs in finite time. Our initial state, which contains two vorticity nulls with a velocity null in between, is of geometrical interest because nulls are a source of separatrices and singularities tend to occur near separatrices.

The finite-time vortex singularity in our model occurs at the stagnation point of the flow. The singularity develops at a point and is locally self-similar. The inviscid flow in the vicinity of the singular point has a complex spatial structure involving two disparate small scales.

There is a spiral structure in the initial conditions that is not essential to the formation of the vortex singularity. The presence of even a small but finite viscosity thwarts the formation of the vortex singularity and the spirals then provide a mechanism for energy cascade from the large to the small spatial scales, as originally envisioned by Lundgren [32]. Once the connection with Lundgren's model is established, the Kolmogorov spectrum follows after time averaging over the life of a vortex tube. As explained by Gilbert [34], the $k^{-5/3}$ spectrum is a robust consequence of time averaging that causes some remarkable cancellations in Lundgren's model.

An important question is how the results of this model can be connected to numerical (as well as real) experiments. There is an extensive database on vortex simulations that has grown out of careful numerical work over the past decade [5–9], including studies of vortex reconnection with antiparallel and orthogonal vortex tubes [48–54] in which no conclusive identification of vortex singularities have yet been made. There are, however, three recent numerical studies [45,46,55] in which finite-time vortex singularities growing as inverse time have been reported. We have made a qualitative connection with the singularities reported in one of them (Ref. [45]).

The model we have developed is quite simple and its assumptions need to be checked by further analytical and numerical studies. Yet the model seems to contain some essential attributes of turbulence. The qualitative picture that emerges from this paper begins with the (tendency of) unbounded local growth of vorticity in finite time at separatrices, defined by a network of nulls that act as attractors for the flow. This process leads to the growth of inviscid small scales. The vorticity intensification is eventually halted in the presence of even a small but finite viscosity. Spiral structures then provide a mechanism by which the Kolmogorov spectrum may be realized.

ACKNOWLEDGMENTS

We are pleased to acknowledge helpful discussions with R. Caflisch, J. Greene, and R. Pelz. We are indebted to S. Tanveer for prompting us to undertake the analysis in Sec. V and to him and S. Kida for constructive criticism of an earlier version of our work. This research was supported by the Air Force Office of Scientific Research Grant No. F49620-93-1-0071 and the National Science Foundation Grant No. ATM-9310157.

[1] A. N. Kolmogorov, C. R. Akad. Sci. SSSR **30**, 301 (1941).
[2] W. Wolibner, Math. Z. **37**, 698 (1933).
[3] V. I. Judovic, USSR Comput. Math. Phys. **3**, 1407 (1963).
[4] T. Kato, Arch. Rat. Mech. Anal. **25**, 188 (1967).
[5] M. Brachet, D. Meiron, B. Nickel, S. Orszag, and U. Frisch, J. Fluid Mech. **130**, 411 (1983).
[6] E. D. Siggia, Phys. Fluids **28**, 794 (1985).
[7] C. R. Anderson and C. Greengard, Commun. Pure Appl. Math. **42**, 1123 (1989).
[8] R. M. Kerr and F. Hussian, Physica D **37**, 474 (1989).
[9] A. Pumir and E. D. Siggia, Phys. Fluids A **2**, 220 (1990).
[10] R. Grauer and T. C. Sideris, Phys. Rev. Lett. **67**, 3511 (1991).
[11] A. Pumir and E. Siggia, Phys. Fluids A **4**, 1472 (1992).
[12] X. Wang and A. Bhattacharjee, in *Topological Aspects of the Dynamics of Fluids and Plasmas,* edited by H. K. Moffatt, G. M. Zaslavsky, M. Tabor, and P. Comte (Kluwer Academic, Dordrecht, 1992), pp. 303–308.
[13] S. Childress, Phys. Fluids **30**, 944 (1987).
[14] W. E and C.-W. Shu, Phys. Fluids **6**, 49 (1994).
[15] W. W. Grabowski and T. L. Clark, J. Atmos. Sci. **48**, 527 (1991).
[16] R. Caflisch, Physica D **67**, 1 (1993).
[17] M. D. Kruskal and R. M. Kulsrud, Phys. Fluids **1**, 265 (1958).
[18] V. Arnol'd (unpublished).
[19] H. K. Moffatt, J. Fluid Mech. **159**, 359 (1985).
[20] H. K. Moffatt, J. Fluid Mech. **166**, 359 (1986).
[21] C. C. Hegna and A. Bhattacharjee, in *Topological Fluid Mechanics,* edited by H. K. Moffatt and A. Tsinober (Cambridge University Press, Cambridge, 1990), pp. 206–215.
[22] C. C. Hegna and A. Bhattacharjee, J. Fluid Mech. **227**, 527 (1991).
[23] J. M. Greene, Phys. Fluids B **5**, 2355 (1993).
[24] A. Lifschitz and E. Hameiri, Phys. Fluids A **3**, 2644 (1991).
[25] S. Friedlander and M. Vishik, Phys. Rev. Lett. **66**, 2204 (1991).
[26] J. M. Greene, J. Geophys. Res. **93**, 8583 (1988).
[27] Y.-T. Lau and J. M. Finn, Astrophys. J. **350**, 672 (1990).
[28] A. Bhattacharjee and X. Wang, Phys. Rev. Lett. **69**, 2196 (1992).
[29] S. Tanveer and C. G. Speziale, Phys. Fluids A **5**, 1456 (1993).
[30] P. Constantin, SIAM Rev. **36**, 73 (1994).
[31] P. Constantin and C. Fefferman, Indiana Univ. Math. J. **42**, 775 (1994).
[32] T. S. Lundgren, Phys. Fluids **25**, 2193 (1982); Phys. Fluids A **5**, 1472 (1993).
[33] D. I. Pullin and P. G. Saffman, Phys. Fluids A **5**, 126 (1992).
[34] A. D. Gilbert, Phys. Fluids A **5**, 2831 (1993).
[35] S. W. H. Cowley, Radio Sci. **8**, 903 (1973).
[36] S. Fukao, U. Masayuki, and T. Takao, Rep. Ionos. Space Res. Jpn. **29**, 133 (1975).
[37] S. Kida, J. Phys. Soc. Jpn. **54**, 2132 (1985).
[38] J. Leorat, Ph.D. thesis, Universite de Paris, 1975 (unpublished) pp. 125–129.
[39] P. Vieillefosse, J. Phys. (Paris) **43**, 837 (1982).
[40] B. J. Cantwell, Phys. Fluids A **4**, 782 (1992).
[41] G. K. Batchelor, *An Introduction to Fluid Mechanics* (Cambridge University Press, Cambridge, 1967), pp. 92–93.
[42] S. Childress, G. R. Ierly, E. A. Spiegel, and W. R. Young, J. Fluid Mech. **203**, 1 (1989).
[43] J. T. Stuart, in *Symposium to Honor C. C. Lin,* edited by D. J. Benney, F. H. Shu, and C. Yuan (World Scientific, Singapore, 1987).
[44] T. Kato, J. T. Beale, and A. Majda, Commun. Math. Phys. **94**, 61 (1989).
[45] R. M. Kerr, Phys. Fluids A **5**, 1725 (1993).
[46] O. N. Boratav and R. B. Pelz, Phys. Fluids A **6**, 2757 (1994).
[47] R. Caflisch (private communication).
[48] W. T. Ashurst and D. I. Meiron, Phys. Rev. Lett. **58**, 1636 (1987).
[49] S. Kida and M. Takaoka, Phys. Fluids **30**, 2911 (1987).
[50] A. Pumir and R. M. Kerr, Phys. Rev. Lett. **58**, 1632 (1987).
[51] M. V. Melander and N. J. Zabusky, Fluid Dyn. Res. **3**, 247 (1988).
[52] M. V. Melander and F. Hussain, Phys. Fluids A **1**, 633 (1989).
[53] M. J. Shelley, D. E. Meiron, and S. A. Orszag, J. Fluid Mech. **246**, 613 (1993).
[54] O. N. Boratov, R. B. Pelz, and N. J. Zabusky, Phys. Fluids A **4**, 581 (1992).
[55] J. B. Bell and D. L. Marcus, Commun. Math. Phys. **147**, 371 (1992).

Sufficient condition for a finite-time singularity in a high-symmetry Euler flow: Analysis and statistics

C. S. Ng and A. Bhattacharjee
Department of Physics and Astronomy, The University of Iowa, Iowa City, Iowa 52242
(Received 20 March 1995)

A sufficient condition is obtained for the development of a finite-time singularity in a highly symmetric Euler flow, first proposed by Kida [J. Phys. Soc. Jpn. **54**, 2132 (1995)] and recently simulated by Boratav and Pelz [Phys. Fluids **6**, 2757 (1994)]. It is shown that if the second-order spatial derivative of the pressure (p_{xx}) is positive following a Lagrangian element (on the x axis), then a finite-time singularity must occur. Under some assumptions, this Lagrangian sufficient condition can be reduced to an Eulerian sufficient condition which requires that the fourth-order spatial derivative of the pressure (p_{xxxx}) at the origin be positive for all times leading up to the singularity. Analytical as well as direct numerical evaluation over a large ensemble of initial conditions demonstrate that for fixed total energy, p_{xxxx} is predominantly positive with the average value growing with the numbers of modes. [S1063-651X(96)13008-7]

PACS number(s): 47.27.Cn

The mechanisms by which a fluid generates intense small-scale dynamics are crucial to our understanding of turbulence. Once small scales are created spontaneously, dissipation intervenes, and the dynamical balance between the two processes determines the character of turbulence. In particular, the dissipation rate of Navier-Stokes turbulence depends crucially on how the vorticity scales with the Reynolds number. Therefore, it is of great importance to study how small scales can be generated in a fluid by the action of vortex stretching controlled by the nonlinearities in the three-dimensional (3D) Euler equation (which is the infinite-Reynolds-number limit of the Navier-Stokes equation),

$$\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v}\cdot\nabla\mathbf{v} = -\nabla p. \tag{1}$$

Here, for incompressible velocity fields, the self-consistent pressure p must satisfy the equation $\nabla^2 p = -\nabla\cdot(\mathbf{v}\cdot\nabla\mathbf{v})$. The main question is whether the solution to Eq. (1) becomes singular in finite time for a smooth initial condition with finite energy.

Mathematicians have provided some useful and rigorous constraints on the nature of possible singularities in 3D flows [1–4], but a physical model which explicitly demonstrates the singularity in a mathematically rigorous way remains elusive. It has been claimed [5,6] that a recent analytical model developed for a symmetric initial condition exhibits a finite-time singularity, but the demonstration relies on some strong assumptions which, while physically plausible, have not yet been substantiated formally or verified by a suitably designed numerical experiment. The analytical results [5] and [6] are suggestive: if one begins from an initial state with symmetries that are preserved by the Euler equation for all time, then the problem of finite-time singularities of the Euler equation could be somewhat more tractable.

This paper is stimulated by the recent numerical experiment of Boratav and Pelz (BP) [7] on a highly symmetric initial flow field, first proposed by Kida [8]. Due to the high symmetry of the Kida flow, BP were able to simulate the 3D Navier-Stokes equation with high spatial resolution and relatively small viscosity. (For the computer runs reported by BP, the Reynolds number Re ($=1/\nu$) varies from 1000 to 5000 and the maximum total resolution is 1024^3.) Within the limits of the spatial resolution, BP report that the maximum vorticity scales as $(t_c - t)^{-1}$, and attribute its eventual saturation to the presence of viscosity. In a subsequent paper [9], BP report a loss of regularity in the strain tensor, and find, furthermore, that the spatial locations of the almost-divergent strain and vorticity are not coincident.

The principal goal of this paper is to present a sufficient condition for the development of a finite-time singularity in the Kida flow. The demonstration of this sufficient condition provides physical insight into a possible mechanism for singularity formation in this highly symmetric geometry. Two forms of this sufficient condition are given: a Lagrangian form for a moving point, and a more useful Eulerian form for a stationary point (the origin) that can be derived from the Lagrangian form under some assumptions. Though we are unable to provide an analytical proof, there is some numerical evidence [10] that the Eulerian sufficient condition is satisfied for the specific initial condition used in the numerical experiment of BP. We present additional statistical evidence that the underlying symmetries of the Kida flow make it highly probable that this condition is also valid for a large ensemble of initial conditions.

The symmetries of Kida flows have been discussed in detail in [8]. Here we build these symmetries into the representations for $\mathbf{v}$ and p. The components of the velocity field $\mathbf{v}=(v_x,v_y,v_z)$ can be written as $v_x = u(x,y,z)$, $v_y = u(y,z,x)$, $v_z = u(z,x,y)$, where u can be expressed in Fourier series,

$$u(x,y,z) = \sum_{lmn} a_{lmn} \sin lx \cos my \cos nz. \tag{2}$$

Here (l,m,n) are natural numbers which represent the three components of a wave vector ($l \neq 0$). In order to satisfy the symmetries and the condition $\nabla\cdot\mathbf{v}=0$, the following conditions must hold:

 © 1996 The American Physical Society

$$l,m,n \text{ must be all odd or all even,} \tag{3}$$

$$a_{lmn}=(-1)^l a_{lnm}, \tag{4}$$

$$\sum_{lmn}^{C} l a_{lmn}=0, \tag{5}$$

where the last summation (denoted by C) is over all permutations of any three natural numbers (l,m,n), i.e., $la_{lmn}+ma_{mnl}+na_{nlm}=0$. By (2) and (5), it can be seen that for $\mathbf{x}$ close to the origin, $\mathbf{v}=O(|\mathbf{x}|^3)$. In particular, the initial state considered by both Kida [8] and BP is u_0: $a_{1,3,1}=1$, $a_{1,1,3}=-1$, with all other terms set to zero. For this initial state, $v_x=\partial v_x/\partial x=0$ at $t=0$ for all x.

With u represented by (2), it can be shown that the pressure p is of the form $p=\Sigma_{lmn}p_{lmn}\cos lx\cos my\cos nz$, where p_{lmn}, using the Poisson equation for the pressure, is given by

$$p_{lmn}=(lA_{lmn}+mA_{mnl}+nA_{nlm})/(l^2+m^2+n^2).$$

with A_{lmn} defined by

$$(\mathbf{v}\cdot\nabla\mathbf{v})_x\equiv\sum_{lmn}A_{lmn}\sin lx\cos my\cos nz.$$

It can be shown that A_{lmn} also satisfies (3) and (4) (with a_{lmn} replaced by A_{lmn}) as well as (5), provided the summation in (5) is carried out over all (l,m,n). Note that A_{lmn} is a quadratic function of a_{lmn}. If there are two terms a_{lmn} and a_{pqr} in u, then there are terms with the following (l,m,n) values in A_{lmn}:

$$(|l\pm p|,|m\pm q|,|n\pm r|),(|l\pm q|,|m\pm r|,|n\pm p|),$$
$$(l\pm r|,|m\pm p|,|n\pm q|), \tag{6}$$

and their permutations. By (5), we obtain $p_{lmn}=p_{mnl}=p_{nlm}=(-1)^l p_{lnm}$. From (1), we obtain the time evolution equation of a_{lmn},

$$\dot{a}_{lmn}+A_{lmn}-lp_{lmn}=0, \tag{7}$$

where the overdot denotes time derivative. Equation (7) and the mode-generation scheme (6) provide a prescription for the dynamical excitation of modes with increasingly large wave numbers, or a cascade of energy to small scales. If this process happens fast enough, then there may be a finite-time singularity.

Since the Euler equation preserves the Kida symmetries for all time, the selection rules imposed on a_{lmn}, A_{lmn}, and p_{lmn} by these symmetries are preserved by Eq. (7). By (5) and (7), we obtain the useful relation

$$p_{xx}\equiv-\sum_{lmn}l^2p_{lmn}=0, \tag{8}$$

where p_{xx} denotes the second spatial derivative of p at the origin. It follows that $\nabla^2 p=0$ at the origin.

Let us now consider the flow along the line $y=z=0$. By (2), we obtain $v_y=v_z=0$, and $v_x=u=\Sigma_{lmn}a_{lmn}\sin lx$. Note that the vorticity is also identically zero along this line. We define $\alpha\equiv\partial_x v_x$, $\beta\equiv\partial_y v_y$, and $\gamma\equiv\partial_z v_z$, where $\alpha+\beta+\gamma=0$ by the incompressibility condition. The dynamical equation for α can be found by taking the x derivative of (1). We obtain

$$\dot{\alpha}+\alpha^2=-p_{xx}=\sum_{lmn}l^2p_{lmn}\cos lx(t). \tag{9}$$

Equations similar to (9) can be written for β and γ. From (9), we obtain the following sufficient condition for a finite-time singularity: *If $p_{xx}>0$ for all time following a Lagrangian element, then α will be singular in finite time.* However, in order to test this condition, we need to evaluate p_{xx} by following a fluid element. It may be possible to test this condition numerically, but it is not convenient to do so analytically. We, therefore, attempt to obtain an Eulerian sufficient condition that can be evaluated at a fixed point (the origin).

Using (1), we write

$$\dot{v}_x=-p_x=\sum_{lmn}lp_{lmn}\sin lx(t), \tag{10}$$

where the overdot denotes total time derivative along a fluid element moving in a trajectory $x=x(t)$. By (8), for x close to zero, we obtain the Taylor expansion

$$\dot{v}_x=-\frac{p_{xxxx}}{6}x^3+O(x^5), \tag{11}$$

where p_{xxxx} is the fourth-order spatial derivative at the origin, given by $p_{xxxx}=\Sigma_{lmn}l^4p_{lmn}$. Note that at the origin, although $\nabla^2\nabla^2 p=3p_{xxxx}+6p_{xxyy}=0$, p_{xxxx} is nonzero in general. For further reference, note also that $p_{xxxx}(x)$ is a symmetric function of x. The fourth-order derivative p_{xxxx} plays an important role because all other lower-order derivatives vanish at the origin in this highly symmetric flow. From the exact equation (9), we obtain

$$\dot{\alpha}+\alpha^2=-\frac{p_{xxxx}}{2}x^2(t)+O(x^4). \tag{12}$$

It is easy to see, using the selection rules for p_{lmn}, that $p_{xx}(x)$ is symmetric and $p_x(x)$ is antisymetric about $x=\pi/2$. Also, since $p_x=0$ at $x=0$ and $x=\pi/2$, we infer that p_{xx} has to assume both positive and negative values within the range $0<x<\pi/2$. Hence, there must be a region of x in which amplification of α^2 occurs. This leads to the next question: is there always a fluid element in the region of amplification? To answer this question, let us assume that there exists a range $0<x<X(t)$ in which $p_{xxx}(x)>0$ for all time before a possible singularity appears. Furthermore, we assume that $X(t)>C$, where C is a finite positive constant for all time (including $t\to t_c$). Since $p_{xxxx}(0)=p_x(0)=0$ by the symmetry conditions and $p_{xx}(0)=0$ by (8), it follows from the assumption above and by simple integration from $x=0$ that the quantities $p_x(x)$, $p_{xx}(x)$, and $p_{xxx}(x)$ are also positive within the range $0<x<X$. Then by (10) and the fact that $v_x(x,t=0)=0$, there exists a fluid element with the Lagrangian coordinate $x(t)$ within this range $(0,C)$ always accelerating towards the origin $x=0$. However, since the condition $v_x(x=0,t)=0$ is always maintained by the symmetry of the flow, the fluid element cannot pass through the origin.

Therefore, $x(t)$ always decreases but remains positive, even when the velocity becomes very large (or even singular). This seeming contradiction between what the trajectory tends to do and what it is constrained to do by symmetry is precisely the mechanism for the development of the finite-time singularity. On the one hand, $x(t)$ is always accelerated towards the origin and tends to reach the origin in finite time. On the other hand, $x(t)$ cannot actually reach the origin because the symmetry conditions forbid it. The system resolves this contradiction by having the velocity derivative $\alpha = \partial v_x/\partial x$ blow up in finite time, since the fluid element with finite and increasing velocity is forced to go infinitesimally close to the point with zero velocity ($x=0$). This behavior is reflected in Eq. (9) according to which α tends to negative infinity in finite time due to the presence of the α^2 term. If the time dependence of α is determined dominantly by the α^2 term, then $\alpha \to (t_c - t)^{-1}$ as $t \to t_c$. Under the assumptions discussed above, we have thus demonstrated that the condition $p_{xxxx} > 0$ at the origin is a sufficient condition for a finite-time singularity.

We caution that the assumption of existence of a finite $X(t)$ is a strong one, and may limit the applicability of our sufficient condition at a fixed point. In some physical cases, as the singularity develops, $X(t)$ may actually tend to zero as $t \to t_c$. If that occurs and a fluid element falls out of the amplification region, t_c may tend to infinity. However, we speculate that if a shrinking $X(t)$ is accompanied by p_{xxxx} growing sufficiently fast, then the finite-time singularity may be supported because a fluid element is then accelerated fast enough to remain in $X(t)$ even as $X(t) \to 0$.

The sufficient condition for singularity (in its moving-point or fixed-point form) does not violate the theorems proved by Beale-Kato-Majda [1] and Ponce [2] which can be essentially summarized as follows: If there occurs a finite-time singularity in an initially smooth Euler flow of finite energy, then the time integral of the maximum norm of the vorticity [1] (deformation tensor [2]) must tend to infinity as $t \to t_c$. We remark that our discussion of the sufficient condition involves the $y=z=0$ axis on which the vorticity is identically zero by symmetry, but there is no restriction on the vorticity off the axis. If the sufficient condition is satisfied, the deformation tensor must be singular near the origin at least as fast as $1/(t_c - t)$, but this leaves open the possibility that the vorticity can blow up at another spatial location. Indeed, BP [9] report that the locations of near-divergent strain and vorticity do not coincide in space. Such a possibility is not inconsistent with [1] and [2]. BP [10] have recently checked that sufficient condition in their numerical simulation for the initial condition in one of their runs (run D3), and their data indicates that p_{xxxx} does remain positive and growing for all times from $t=0$ to the singularity time. (In particular, p_{xxxx} is found to take the following sequence of values: 40.53 at $t=0$, 226.4 at $t=1.5$, 2.642×10^4 at $t=2.0$, 2.802×10^5 at $t=2.125$, and 1.111×10^7 at $t=2.25$. The extrapolated singularity time t_c reported by BP for this run is 2.21.)

Though the numerical evidence presented above is suggestive, it cannot be regarded as definitive proof of the existence of the singularity. Furthermore, we cannot deduce generic properties of Kida flows from the numerical evidence for one specific initial condition. We now proceed to give a statistical demonstration that the positivity of p_{xxxx} is highly probable over a large ensemble of initial conditions.

First, we introduce a minimal, independent set of modes u_n of the Kida flow such that any flow u satisfying the symmetries of the Kida flow can be written as $u = \Sigma_n a_n u_n$, where a_n are real constants. Then, the following is the only possible choice for the set of independent modes that yield a minimum number of terms for each mode and satisfy the symmetries (4) and (5). For any three distinct, odd positive integers l,m,n, there are, in general, two independent modes:

$$\begin{aligned} u_1: &\quad a_{lmn} = -a_{lnm} = m, \quad a_{mnl} = -a_{mln} = -l, \\ u_2: &\quad a_{lmn} = -a_{lnm} = n, \quad a_{nlm} = -a_{nml} = -l. \end{aligned} \tag{13}$$

By (5), the third mode (u_3: $a_{mnl} = -a_{mln} = -n$, $a_{nlm} = -a_{nml} = m$) is not independent and can be written $u_3 = (nu_1 - mu_2)/l$. In practice, we can choose any two of the above set of three as independent. However, if two of the three integers are equal, there is only one independent mode, $a_{lml} = -a_{llm} = 1$. Also, there is no mode for $l=m=n$. The initial condition used by BP, that is, u_0: $a_{1,3,1} = 1$, $a_{1,1,3} = -1$, is clearly an independent mode. For any three even natural numbers, there are also, in general, only two independent modes:

$$\begin{aligned} u_1: &\quad a_{lmn} = a_{lnm} = m, \quad a_{mnl} = a_{mln} = -l, \\ u_2: &\quad a_{lmn} = a_{lnm} = n, \quad a_{nlm} = a_{nml} = -l. \end{aligned} \tag{14}$$

The third mode, which is not independent, is again $u_3 = (nu_1 - mu_2)/l$. If two of the three numbers are equal or one of them is zero, then there is only one independent mode. Also, there is no mode if two numbers are zero or one number is zero and the other two equal, or three numbers are equal.

The quantity p_{xxxx} (at the origin) is a quadratic function of u. We write $p_{xxxx} = P(u,u)$. With the representation of u, we obtain

$$p_{xxxx} = \sum_n a_n^2 P(u_n, u_n) + \sum_n \sum_{m \neq n} a_n a_m P(u_n, u_m). \tag{15}$$

Defining $P_{nm} = [P(u_n, u_m) + P(u_m, u_n)]/2$, we can also write $p_{xxxx} = \Sigma_n \Sigma_m a_n P_{nm} a_m$. We have a constructive proof using MATHEMATICA that $P(u_n, u_n) > 0$. The symbolic manipulations are too long to be presented here and will be discussed in a separate publication. In order to determine the positivity of p_{xxxx}, we now need to consider the second term on the right-hand side of (15), which involves the cross terms. The contributions from the cross terms cannot be neglected in principle, and they can be positive or negative depending on the sign of a_n. To assess their importance, let us consider an example with two modes, i.e., $u = a_m u_m + a_n u_n$, so that

$$P(u,u) = a_m^2 P_{mm} + a_n^2 P_{nn} + 2 a_m a_n P_{mn}. \tag{16}$$

The cross term P_{nm} is given by the relation $2P_{nm} = P(u_n + u_m, u_n + u_m) - P_{nn} - P_{mm}$. Note that the cross term between an odd mode and an even mode is always zero due to

TABLE I. Some results of the Monte Carlo calculations of p_{xxxx} with $\nu=3$ showing that $p_{xxxx}>0$ in most cases. See text and Eqs. (27) and (28) for definitions of the variables.

N	k_N^2	M (units of 10^6)	η (%)	$\langle P\rangle$	$\langle P_D\rangle$	$\langle d\rangle$	$\langle d_D\rangle$	$\langle r\rangle$
11	44	2	1.09	20.4	20.4	21.9	8.6	1.00
37	99	2	0.96	41.9	41.9	45.5	11	1.00
64	136	3	0.82	61.8	61.2	67.6	13	1.01
136	219	3	0.77	98.2	98.0	108	14	1.01
211	296	4	0.70	130	131	141	16	1.00
290	360	7	0.64	161	162	174	17	1.00
449	480	10	0.64	212	214	230	18	0.99
612	587	20	0.61	267	264	290	19	1.01
777	691	30	0.61	307	308	334	20	1.00
945	780	40	0.61	349	351	377	21	0.99
1114	875	50	0.58	394	390	423	22	1.01
1283	963	60	0.56	429	430	464	22	1.00
1455	1043	80	0.58	470	467	505	23	1.01
1624	1123	100	0.59	500	502	537	23	1.00
1973	1275	160	0.54	573	571	616	24	1.00
2144	1352	200	0.54	605	602	675	25	1.01

the selection rules on the pressure. For the positivity of $P(u,u)$ in (16), it is sufficient to have $P_{mm}P_{nn}\geqslant P_{nm}^2$. This relation is found to be true numerically if either u_m or u_n is the initial flow u_0 defined above, or the mode with $a_{3,3,1}=1$ and $a_{3,1,3}=-1$. However, the condition $P_{mm}P_{nn}\geqslant P_{nm}^2$ is not always true for any two modes. Instead, we have shown by using MATHEMATICA that

$$\frac{P_{nm}^2}{P_{mn}P_{nn}}\to 0 \quad \text{as} \quad \frac{k_m^2}{k_n^2}\to 0, \tag{17}$$

where k_m, k_n are the wave numbers of the modes u_m, u_n. The limit (17) is attained faster by odd modes than by the even modes. Hence, for those cases, which make up most of the pairs of modes for a flow spanning wave numbers over several orders of magnitude, the contribution from off-diagonal terms is much smaller than that from diagonal terms.

To estimate the contributions from the cross terms systematically in a general case, let us consider a flow represented by N modes ($n=1,N$) that includes all modes with $k_n<k_N$. For normalization, we define an energy functional

$$E(u)\equiv\frac{1}{\pi^3}\int_0^{2\pi}dx\int_0^{2\pi}dy\int_0^{2\pi}dz u^2=2\sum_{lm}a_{lm0}^2+\sum_{lmn}a_{lmn}^2, \tag{18}$$

which is conserved in an Euler flow. For the specific initial condition used in [7] and [8], we obtain $E(u_0)=2$. Hence, we normalize the u_n in (18) such that $E(u_n)=2$. We now perform a Monte Carlo calculation of $P(u,u)$. The calculation is carried out M times with $M\gg 1$, with a_n chosen randomly each time within a range specified by an energy spectrum in k space of the form $E(k)\propto k^{-\mu}$ to simulate the energy distribution over different length scales. In most of our calculations, we choose $\mu=3$ which is the spectrum observed by BP near $t=t_c$ [7]. (As discussed later, the qualitative trends observed are not sensitive to variations in μ.) In Table I, we report numerical results with $N\leqslant 2144$, or $k_N^2\leqslant 1352$. We define η as the percentage of cases with negative vales for p_{xxxx}, and the averaged quantities are

$$\begin{aligned}\langle P\rangle&\equiv\frac{1}{M}\sum_{i=1}^{M}p_{xxxx},\\ \langle P_D\rangle&\equiv\frac{1}{M}\sum_{i=1}^{M}\sum_{n=1}^{N}a_n^2P_{nn}=\frac{1}{M}\sum_{i=1}^{M}P_D,\end{aligned} \tag{19}$$

where P_D is the contribution from the diagonal terms. Also, we define the deviations and the average ratio as

$$\begin{aligned}\langle d\rangle&\equiv\left[\frac{1}{M}\sum_{i=1}^{M}(p_{xxxx}-\langle P\rangle)^2\right]^{1/2},\\ \langle d_D\rangle&\equiv\left[\frac{1}{M}\sum_{i=1}^{M}(P_D-\langle P_D\rangle)^2\right]^{1/2},\\ \langle r\rangle&\equiv\frac{1}{M}\sum_{i=1}^{M}\frac{p_{xxxx}}{P_D}.\end{aligned} \tag{20}$$

The convergence of the data is tested for large M in two cases. The fluctuations in all quantities are found to be typically less than 2%.

We examine the sensitivity of the results to the assumed form of the energy spectrum by recalculating the $N=177$ case with $\mu=-2,0,2,4,6$. It is found for these cases that η becomes 1.4%, 1.31%, 0.91%, 0.35%, 0.26%, respectively. Hence, the effect of μ is seen not to be qualitatively important. From the data in Table I, we see that η is less than 1% for most cases, except the cases with very small N. If this trends continues to hold for larger N values, then the probability that $p_{xxxx}>0$ is much larger than the probability that $p_{xxxx}<0$. Note also that for all the cases discussed above, $\langle P\rangle\approx\langle P_D\rangle\approx\langle d\rangle$, $\langle r\rangle\approx 1$. This implies that the average con-

tribution to p_{xxxx} is dominantly from the self terms, and that the cross terms mostly cancel each other upon summation. We also see from Table I that $\langle P \rangle \propto k_N^2$ which, in light of our remarks above, is further evidence in support of a finite-time singularity because the data clearly shows the growth of p_{xxxx} as length scales decrease.

The numerical results presented above are based mostly on statistics. Since the deterministic dynamics of Kida flows do not have to follow the most probable path, we cannot regard the evidence above as a dynamical proof that $p_{xxxx} > 0$, However, the evidence does suggest that the condition $p_{xxxx} > 0$ is highly probable and is strongly favored by the symmetry properties of the Kida flow, independent of the precise dynamical details emerging from a specific initial condition.

The high symmetry of the Kida flow enables us to obtain some analytical and numerical results that provide strong physical evidence in support of a finite-time singularity in this class of Euler flows. The assumed symmetry properties preserve the geometric structure of the initial state for all times. Such a singularity may be unstable if the symmetry conditions are relaxed, and so the qualitative implications of these results for more general 3D configurations remain unclear. Till 1990, numerical results on 3D flows were inconclusive despite the sophistication of the numerical methods employed [11–15]. Finite-time singularities have been reported in axisymmetric flows with swirl [16–18], but the results are controversial [19,20]. More recently, prior to the work of BP, two other numerical experiments [21,22] have presented evidence in support of a finite-time singularity in 3D Euler flows. In particular, Kerr's simulation [22] involves antiparallel vortex tubes, and has qualitatively similarities with that of BP in that the singularity occurs in the vicinity of the symmetry axis. The work of Kerr has its antecendents in earlier studies of vortex reconnection with antiparallel and orthogonal vortex tubes [23–29].

In conclusion, we have proposed a sufficient condition for a finite-time singularity in a Kida flow. We have shown that if the second-order spatial derivative of the pressure (p_{xx}) is positive following a Lagrangian element (on the x axis), then a finite-time singularity must occur. Under some assumptions, this Lagrangian sufficient condition can be reduced to an Eulerian sufficient condition which requires that the fourth-order spatial derivative of the pressure (p_{xxxx}) at the origin be positive for all times leading up to the singularity. Though we are unable to provide an analytical proof that this is indeed satisfied in the simulation of BP [7], there is numerical evidence [10] in support of this condition. Furthermore, we have presented strong physical evidence which suggests that it is highly probable that the Eulieran form of the sufficient condition for singularity is satisfied for a large ensemble of initial conditions.

We are grateful to Dr. O. Boratav and Professor R. Pelz for communicating their numerical data on the sufficient condition for singularity. This research is supported by the NSF Grant No. ATM 93-10157 and the AFOSR Grant No. F49620-93-1-0071.

[1] J. T. Beale, T. Kato, and A. Majda, Commun. Math. Phys. **94**, 61 (1984).
[2] G. Ponce, Commun. Math. Phys. **98**, 349 (1985).
[3] P. Constantin, SIAM Review **36**, 73 (1994).
[4] P. Constantin and C. Fefferman, Indiana Univ. Math. J. **42**, 775 (1994).
[5] A. Bhattacharjee and X. Wang, Phys. Rev. Lett. **69**, 2196 (1992).
[6] A. Bhattacharjee, C. S. Ng, and X. Wang, Phys. Rev. E **52**, 5110 (1995).
[7] O. N. Boratav and R. B. Pelz, Phys. Fluids **6**, 2757 (1994).
[8] S. Kida, J. Phys. Soc. Jpn. **54**, 2132 (1985); S. Kida and Y. Murakami, *ibid.* **55**, 9 (1986).
[9] O. N. Boratav and R. B. Pelz, Phys. Fluids **7**, 895 (1995).
[10] O. N. Boratav and R. B. Pelz (private communication).
[11] M. Brachet, D. Meiron, B. Nickel, S. Orszag, and U. Frisch, J. Fluid Mech. **130**, 411 (1983).
[12] E. D. Siggia, Phys. Fluids **28**, 794 (1985).
[13] C. R. Anderson and C. Greengard, Comm. Pure App. Math. **42**, 1123 (1989).
[14] R. M. Kerr and F. Hussain, Physica D **37**, 474 (1989).
[15] A. Pumir and E. D. Siggia, Phys. Fluids A **2**, 220 (1990).
[16] R. Grauer and T. C. Sideris, Phys. Rev. Lett. **67**, 3511 (1991).
[17] A. Pumir and E. Siggia, Phys. Fluids A **4**, 1472 (1992).
[18] R. Caflisch, Physica D **67**, 1 (1993).
[19] X. Wang and A. Bhattacharjee, in *Topological Aspects of the Dynamics of Fluids and Plasmas*, edited by H. K. Moffatt, G. M. Zaslavsky, M. Tabor, and P. Comte (Kluwer Academic, Dordrecht, The Netherlands, 1992), pp. 303–308.
[20] W. E. and C.-W. Shu, Phys. Fluids **6**, 49 (1994).
[21] J. B. Bell and D. L. Marcus, Commun. Math. Phys. **147**, 371 (1992).
[22] R. M. Kerr, Phys. Fluids A **5**, 1725 (1993).
[23] W. T. Ashurst and D. I. Meiron, Phys. Rev. Lett. **58**, 1632 (1987).
[24] S. Kida and M. Takaoka, Phys. Fluids **30**, 2911 (1987).
[25] A. Pumir and R. M. Kerr, Phys. Rev. Lett. **58**, 1632 (1987).
[26] M. V. Melander and N. J. Zabusky, Fluid Dyn. Res. **3**, 247 (1988).
[27] M. V. Melander and F. Hussain, Phys. Fluids A **1**, 633 (1989).
[28] M. J. Shelley, D. E. Meiron, and S. A. Orszag, J. Fluid Mech. **246**, 613 (1993).
[29] O. N. Boratav, R. B. Pelz, and N. J. Zabusky, Phys. Fluids A **4**, 581 (1992).